

UNIVERSITY OF HELSINKI

The effects of visual mode, accent and attitude on EFL listening comprehension with authentic material

Tuomas Iivanainen
Master's thesis
English philology
Department of Modern Languages
University of Helsinki
April 2016

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Humanistinen tiedekunta		Laitos – Institution – Department Nykykielten laitos	
Tekijä – Författare – Author Tuomas Iivanainen			
Työn nimi – Arbetets titel – Title The effects of visual mode, accent and attitude on EFL listening comprehension with authentic material			
Oppiaine – Läroämne – Subject Englantilainen filologia			
Työn laji – Arbetets art – Level Pro gradu		Aika – Datum – Month and year Huhtikuu 2016	
		Sivumäärä – Sidoantal – Number of pages 86	
Tiivistelmä – Referat – Abstract			
<p>Tutkimuksen tavoitteena on selvittää, miten visuaalinen moodi, aksentti ja asenne vaikuttavat lukiolaisten englannin kuullun ymmärtämiseen, kun kuuntelussa käytetään autenttista materiaalia. Tutkimuksessa vertaillaan audiovisuaalista kuullun ymmärtämistehtävää sekä tavanomaisempaa kuullun ymmärtämistehtävää, jossa puhujaa ei nähdä. Tutkimusongelmia on neljä: Kuinka puhujan näkeminen vaikuttaa englannin kuullun ymmärtämiseen? Kuinka puhujan aksentti vaikuttaa kuulijan ymmärrykseen? Miten kuulijan asenne puhujan aksenttia kohtaan vaikuttaa ymmärrykseen? Millainen on kolmen tutkitun muuttujan, visuaalisen moodin, aksentin ja asenteen, suhde toisiinsa?</p> <p>Teoreettisena pohjana käytetään aiempia tutkimuksia audiovisuaalisen tekstin ymmärtämisestä, kuulijalle vieraiden aksenttien ymmärtämisestä sekä asenteen vaikutuksesta kuullun ymmärtämiseen. Teoria on koostettu useista tutkimuksista, sillä tälle aiheelle ei ole entuudestaan olemassa kattavaa teoriaa. Tutkimuksessa käytetään monivalintatehtävää mittaamaan ymmärrystä audiovisuaalisessa ja audiaalisessa kontekstissa. Koehenkilöiden asenteita puhujan aksenttia kohtaan mitataan kaksinapaisella arviointikyselyllä. Koeryhmä koostuu 39:stä lukio-opiskelijasta, joista 27 on ensimmäisen vuoden opiskelijoita ja 12 kolmannen vuoden opiskelijoita. Pienen otteensa vuoksi tutkimus on tapaustutkimus.</p> <p>Tutkimustulosten perusteella voidaan todeta, että puhujan näkeminen helpottaa kuullun ymmärtämistä silloin, kun hän elehti niin, että kuulija voi rakentaa merkitystä sekä nähdyn että kuullun perusteella. Huulilta lukemisen ei havaittu hyödyttävän kuullun ymmärtämisessä, kun puhuja puhui kuulijalle vieraalla aksentilla. Vaikka visuaalinen moodi pääsääntöisesti helpottikin kuullun ymmärtämistä, saattoivat monitulkintaiset eleet olla jopa harhaanjohtavia. Kuulijalle vieras aksentti vaikeutti kuullun ymmärtämistä selvästi. Vaikka puhujan näkeminen helpotti kuullun ymmärtämistä, se ei tasoittanut aksenttien välisiä haastavuuseroja. Kuulijan asenteen puhujan aksenttia kohtaan havaittiin vaikuttavan subjektiiviseen käsitykseen ymmärtämisestä sen sijaan, että se olisi vaikuttanut objektiiviseen tuloksista mitattuun ymmärryksen tasoon. Tämä oli yllättävää, sillä aiemmat tutkimukset ovat painottaneet asenteen vaikutusta tärkeänä osana kuullun ymmärtämistä. Koehenkilöiden asenteet puhujaa kohtaan olivat positiivisemmat silloin, kun he näkivät puhujan.</p> <p>Koska tutkimuksessa keskitytään vieraan kielen kuullun ymmärtämiseen, siitä on järkeenkäypää johtaa pedagogisia sovelluksia. Kuulijalle vieraiden aksenttien käyttäminen vaikeuttaa kuullun ymmärtämistä selvästi, mikä tulee ottaa huomioon kuullun ymmärtämistehtäviä suunniteltaessa. Audiovisuaalisia kuullun ymmärtämistehtäviä suunniteltaessa tulee kaikki puhujan eleet tulkita huolellisesti, jotta niiden vaikutus kuullun ymmärtämiseen voidaan ennakoita.</p>			
Avainsanat – Nyckelord – Keywords Kuullun ymmärtäminen, audiovisuaalisuus, aksentit, asenteet, englantia vieraana kielenä			
Säilytyspaikka – Förvaringställe – Where deposited			
Muita tietoja – Övriga uppgifter – Additional information			

Table of contents

1	Introduction.....	1
2	Understanding audiovisual texts.....	4
2.1	The pros and cons of multimedia	4
2.2	The cognitive theory of multimedia learning	5
2.3	Schema theory	8
2.4	Cognitive load and the constraints of working memory	10
2.5	Nonverbal communication	12
2.6	Previous studies on understanding audiovisual texts	13
3	Understanding accents	16
3.1	Prosodic features	16
3.2	Knowledge systems	18
3.3	Coping with accents	18
3.4	The interlanguage speech intelligibility benefit	19
3.5	Previous research on understanding accents	20
4	Attitudes towards accents	22
4.1	Why does attitude matter?.....	22
4.2	The formation of attitudes	23
4.3	Accent attitudes and EFL learners.....	27
4.4	Techniques for testing accent attitudes.....	28
5	Material and methods.....	31
5.1	Audiovisual texts.....	31
5.2	Multiple choice test	33
5.3	Speaker evaluation questionnaire	36
5.4	The sample and conducting the test	37
6	Results and discussion	41
6.1	The effect of the visual mode	41

6.1.1	General effects of the visual mode.....	41
6.1.2	The effect of visual mode on the intelligibility of accents.....	43
6.2	Detailed findings of the visual mode.....	43
6.3	The effect of accent	45
6.3.1	Differences between the intelligibility of accents	46
6.3.2	Interpretation of the mean scores	47
6.3.3	The high intelligibility of the Japanese accent.....	48
6.3.4	The role of speech rate	50
6.3.5	Comparison to previous studies	50
6.4	Discussion of individual questions.....	51
6.4.1	Facilitative visual mode	52
6.4.2	Hindering visual mode	55
6.4.3	Questions with ambiguous results.....	58
6.4.4	Equally difficult conditions.....	60
6.5	Results of the speaker evaluation questionnaire.....	62
6.5.1	General observations on the status and solidarity dimensions.....	62
6.5.2	Intelligibility and comprehensibility	65
6.5.3	The role of attitude	67
6.5.4	Listener irritation.....	68
6.5.5	Estimated speech rate.....	70
6.5.6	Estimated usefulness of video.....	71
7	Summary and conclusions	73
	References	79
	List of the videos used in the MC test	86
	Appendices	87
	Appendix 1. Transcriptions of the texts used in the MC task	87

Appendix 2. The MC test in Finnish	94
Appendix 3. The MC test in English	110
Appendix 4. The SE questionnaire in Finnish	126
Appendix 5. The SE questionnaire in English	127

List of figures

Figure 1. The process of multimedia learning	7
Figure 2. Mean scores in the status and solidarity dimensions	63

List of tables

Table 1. Results of the MC test and the differences between groups	42
Table 2. Results of the MC test displayed as mean scores of the groups	45
Table 3. The mean intelligibility scores subtracted from that of RP	47
Table 4. The actual and estimated usefulness of video ratings	71

1 Introduction

This research explores Finnish upper secondary school English as a foreign language (EFL) students' listening comprehension skills, focusing on three major variables: accent, visual mode and attitude. The research is a case study of two upper secondary school EFL learner groups' performance in a listening comprehension test and their opinions in a Speaker Evaluation questionnaire.

Studying the accent variable is timely and essential since there will be a new Core Curriculum, in which the diversity of English language will be taken into account, for Finnish schooling in 2016. It is stated in the 2016 Core Curriculum that a specific goal of the EFL teaching is to develop the learner as a user and operator of the language in a culturally varied world in local, national, European and global communities (Opetushallitus 2015, 109, my translation). Furthermore, accented speech has already been used in the listening comprehension tests of the Matriculation Examination which is a nationwide test in which the final year students of upper secondary school participate.

As regards the visual mode, this research is topical since in the Finnish media, there has been speculation of the possibility of implementing videos to the Matriculation Examination in the near future as the examination will be updated into electronic form (Kujala & Byholm 2014). The role of attitude is in close connection to testing the effect of accent; according to recent studies, the attitude towards an accent plays a major role in listening comprehension (Pihko 1997, Major et al. 2005).

There have been only a few researches in the area of understanding English accents in Finland although the importance of giving attention to accents in EFL teaching was recognized already in 1995 by Pihko. While studying Finnish EFL learners' understanding of different English accents, she stated that it is reasonable that the goal set for EFL learners is awareness of the great variation of English and understanding at least those native and non-native variants they are most likely to encounter (Pihko 1995, 18).

This research sets out to find out what the current situation in understanding English accents is. As this study encompasses the audiovisual and attitudinal variables as well as the accent related variable, it is the first of its kind. This is because

previously, only the effects of accent and attitude together, accent alone, attitude alone or visual mode alone on listening comprehension have been researched. Thus, this study will give valuable new research results not only in the Finnish context but on the area of studying listening comprehension, as well.

The importance of taking accents into account in EFL teaching in the Finnish context can be demonstrated by Ranta's research results which indicated that 62% of Finnish upper secondary school students' language use outside of school was English as a lingua franca whereas only 25% was English with a native speaker (Ranta 2010, 162). Furthermore, while studying the language use in five globally operating Finland-based companies Louhiala-Salminen and Kankaanranta found out that approximately 70% of English communication in these companies took place between non-native speakers of English (Louhiala-Salminen & Kankaanranta 2012, 266). In the light of these studies, it is reasonable to assume that Finnish students would significantly benefit from being familiarized to different accents of English at school, both native and non-native. As suggested by several researchers, the ability to understand different accents develops only by being exposed to different accents (Gass & Varonis 1984, Pihko 1997, Bradlow & Bent 2003a & 2008, Adank et al. 2009).

As the three major variables that affect listening comprehension are the main focus of this research, the research questions target the potential effects of these variables. The four research questions to which this research tries to find an answer are:

1. How does seeing the speaker affect EFL listening comprehension?
2. How does the speaker's accent affect their intelligibility to EFL learners?
3. How does the listener's attitude towards the speaker's accent affect the intelligibility of the speaker?
4. What is the relationship between the three variables: visual mode, accent and attitude?

The first three questions are quite self-explanatory, but the fourth one requires some elaboration. This research not only aims to find out information on the effects of the three variables in isolation but sets to search for tendencies in how they interact. The goals are to find out if the visual mode has an effect on understanding of accents or

vice versa and if visual mode plays a role in forming an attitude about the speaker. Eyes will be kept open for unexpected interrelationships between the variables, as well; it is hard to predict how they act together since these three variables have not been taken into account simultaneously in previous studies.

The structure of this study follows the three variables studied closely; the visual mode will always be discussed first, then accent and finally the role of attitude. The theory section of this research is three-fold: first the theory of understanding audiovisual texts is discussed, secondly the theory of accent as a variable of listening comprehension is explored and finally, the theory of the effect of attitude on accent on understanding a foreign language is explained. After the theory section, information on the materials and method of the study will be given. Then the results of the listening comprehension test and Speaker Evaluation questionnaire will be presented and discussed. The Results and Discussion section will follow the structure of the theory section: audiovisual aspects will be dealt with first, the effect of accent second and the meaning of attitude third. Finally, in the concluding chapter, the limitations as well as theoretical and pedagogical implications of this study and implications for further research will be discussed.

2 Understanding audiovisual texts

“Despite all technical innovations [...] the acquisition of information through any technical system is subject to the constraints of human information processing.” (Tardieu & Gyselinck 2003, 3)

This section will explain the theory behind understanding audiovisual texts and what kind of skills this comprehension requires. The theoretical aspects introduced in this section are essential in discussing the results of this study as well as in creating a suitable test to elicit research data. As for the quote above, it embodies the relationship between the ambition we have as humans for applying new technology to practical use and our inability to make use of all of the possible applications it provides us because of the limits of our brains. In the case of this research, the new technology is of course the use of video in listening comprehension for language learning purposes.

At this point it is relevant to note that in this study the term ‘text’ refers to a multimodal presentation in its entirety as is conventional in studies of this area. There are various types of multimodal presentations but the type that is in the focus of this research is the audiovisual text. As the term suggests, audiovisual texts consist of both audial and visual modes. Examples of audial modes are music, speech, sound effects, background noise, etc., basically everything that can be perceived by hearing. Examples of visual mode include everything that can be seen: colors, static pictures and motion picture, written text such as subtitles, etc. The word ‘story’ is used to refer to the content of the story the speaker of the text tells.

2.1 The pros and cons of multimedia

Discourse comprehension and communication in general does not rely solely on the understanding of auditory representations but visual processing plays a vital role, as well (Baltova 1994, 508). Multimedia gives language learners several parallel possibilities of developing their comprehension skills; they may simultaneously have sound, text and picture available to them as well as sometimes the possibility to manipulate these things by for example pausing and rewinding a video (Guichon &

McLornan 2008, 86). The obvious benefit of the video is that it can replicate authentic face-to-face communication situations much more closely than audio alone. The visual mode may also provide learners with additional informational sources such as the speaker's body language, physical actions (Guichon & McLornan 2008, 86), movement of lips (Baltova 1994, 509), facial expressions, visible objects referred to and the setting of the communication situation (Baltova 1999, 34). Furthermore, the multiple modalities of video can help the viewer direct their attention and predict what will be talked about next.

However, with the benefits of video also come its disadvantages; additional sources of information impose more demands on the learner's attention and conflicting information can even disperse the learner's attention (Baltova 1994, Guichon & McLornan 2008). Furthermore, because the different modes of a multimedia text may not be understandable in isolation, learners have to divide their attention between multiple sources of information (Tardieu & Gyselinck 2003, 6). Controlling all aspects of the verbal and visual input may require more attention than the learners are actually capable of (Baltova 1994, 510).

2.2 The cognitive theory of multimedia learning

The cognitive theory of multimedia learning is essential for this research, because a test to measure intelligibility of an audiovisual text has to be created in order to elicit the research data. The test must be designed carefully so that it will indeed measure intelligibility and not something else. The contemporary cognitive theory of multimedia learning leans upon three theoretical assumptions: 1) *Dual channels*: humans possess separate channels for processing visual and auditory information, 2) *Limited capacity*: humans are limited in the amount of information that they can process in each channel at one time, 3) *Active processing*: humans engage in active learning by attending to relevant incoming information, organizing selected information into coherent mental representations, and integrating mental representations with other knowledge (Mayer 2001, 44).

Human brain engages in five separate processes during multimedia learning (Mayer 2001, 54). Although presented as a list, these processes do not necessarily occur in this specific order. Actually, they may be simultaneous and can occur many

times during an audiovisual presentation. Pictures and images referred to in the explanations of these five processes do not only mean static images but refer to chunks of animations and videos, as well. The five processes are: 1) selecting relevant words for processing in verbal working memory, 2) selecting relevant images for processing in visual working memory, 3) organizing selected words into a verbal mental model, 4) organizing selected images into a visual mental model, 5) integrating verbal and visual representations as well as prior knowledge (Mayer 2001, 54-57). Figure 1 demonstrates these processes and the next few paragraphs explain how to interpret it.

The first box on the left named *Multimedia Presentation* refers to the audiovisual text in which there are pictures and words as either speech or text. Here too, pictures refer to both static images and video. This is the presentation which the perceiver decodes in their brain.

The second box portrays sensory memory in which information is stored only for a very limited time, just enough for it to be moved into working memory. Information is received by sensory organs, in this case ears and eyes, which recognize portions of this input while it is momentarily stored in the sensory memory (Yi'ian 1998, 24). The arrow from *pictures* to *eyes* indicates that eyes perceive the visual information of the audiovisual text. There are arrows from *words* to both *eyes* and *ears* because words may be presented either as written text or spoken sounds and therefore words can be perceived either by ears or eyes.

The first and second processes or the selection of relevant words for processing in verbal working memory and selection of relevant images for processing in visual working memory take place between the second and third boxes. Because of the limited capacity of working memory, only the most relevant parts of the verbal and visual input are selected for processing (Mayer 2001, 55).

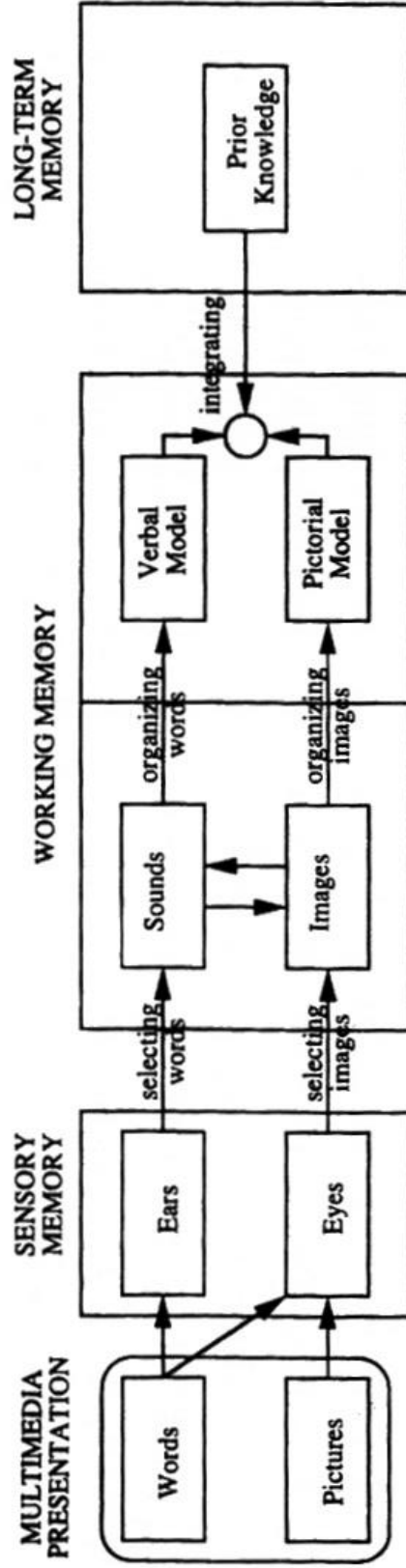


Figure 1. The processes of multimedia learning (Mayer 2001, 44).

There are arrows between *sounds* and *images* in the third box called *Working Memory* because it is possible that some sounds are processed as images and vice versa. This may happen for example if the sound of smacking lips is associated with the visual action of smacking lips. The information that has first been perceived as sound is thus transferred into visual memory.

The third and fourth processes or the organization of selected words into a verbal mental model and the organization of chosen images into a visual mental model take place within the *Working Memory* box. These processes start when a word sound base and a visual picture base have been formed by selecting the relevant sounds and pictures. These bases have to be organized into coherent representations, knowledge structures called *verbal model* and *pictorial model*. This is done by making connections between the selected items. Limited capacity plays a role here, as well, and only simple structures such as a cause-and-effect chains can be structured (Mayer 2001, 56).

The fifth process, the integration of verbal and visual representations as well as prior knowledge is illustrated in the right side of the *Working Memory* box. Relevant prior knowledge is retrieved from *long term memory* and it is integrated with the *verbal and pictorial models*. This integrated product is the understanding of the audiovisual text and it is represented as the small circle to which the arrows from *verbal model*, *pictorial model* and *prior knowledge* point.

2.3 Schema theory

Combined auditory material and visual cues are processed through several levels of prior knowledge; linguistic knowledge, knowledge of the particular context and knowledge of the world (Baltova 1994, 508). Schema theory describes the information processing of human brain and therefore helps us understand the role of prior knowledge in understanding audiovisual texts.

A schema is a complex but organized network, which connects an individual's past experience and knowledge, and it is essential in accessing and comprehending new information (Liang 2015, 2). Schema theory has it that information is retained in long-term memory as schemas that are stable and not easily forgotten. The formation

and development of schemas is an active and dynamic process; schemas are enlarged as the individual gains more knowledge and experience. When the individual receives new information, relevant schemas are stimulated in order to analyze, edit and synthesize the information. When the new information is comprehended, the schemas are renewed and augmented (Liang 2015, 2). Skillful performance develops by building an increasing number of more and more complex schemas by combining elements consisting of lower level schemas into higher level schemas. Schemas may incorporate vast amounts of information if the learning process occurs over a long period of time (Paas & Kester 2006, 281).

According to schema theory there are two ways to handle information: *top-down processing* and *bottom-up processing* (Liang 2015, 2). In top-down processing the integration of new and old information is achieved by processing the language input with the help of prior knowledge. The information stored in schemas is used to predict what is to be presented, to understand what has been stated and to test and adopt the grasped ideas (Liang 2015, 2).

In bottom-up processing the information processing starts from the simplest language units, sounds and syllables. The processing advances to more complicated units of language after the simpler units have been dealt with; if we are to use the decoding of written text as an example, words, sentences, paragraphs, chapters, and the whole text are processed in turns. In the case of audiovisual texts, the boundaries of these hierarchical units are much more obscure. For this reason, written text was used as an example even though it is not the main focus of this study. Bottom-up processing handles schemas relating to language in particular. The more phonetic, lexical, syntactic and contextual knowledge of a language is stored in schemas, the faster and more accurately the language input can be decoded (Liang 2015, 2).

These two types of information processing do not function separately but rather work in unison. While the comprehension of linguistic units can activate background knowledge, the stimulation of background knowledge can lead into deeper understanding of the language input (Liang 2015, 2). The proficiency in the target language influences the harmony between top-down and bottom-up processing: the less knowledgeable the listener is in the language in question, the more they rely on world knowledge or top-down processing. Likewise, the more knowledgeable they are in the target language the more they can rely on the acoustic signal or bottom-up processing (Pihko 1994, 83).

2.4 Cognitive load and the constraints of working memory

Working memory is a system of limited capacity that ensures a double function of dealing with information and temporarily holding information (Tardieu & Gyselinck 2003, 12). The limitations of working memory are lessened for the more experienced learners when dealing with previously learned information that has been stored in schemas, because working memory can treat schemas as single units and retrieve the relevant information with less effort (Paas & Kester 2006, 282). With practice and time, the cognitive processes required to complete a task become more automated, freeing cognitive resources for other activities (Tardieu & Gyselinck 2003, 5). The next few paragraphs explain what the constraints of working memory exactly consist of.

The Cognitive Load Theory attempts to account for the outcomes of the limitations of the human information-processing system. The theory starts with the assumption that working memory has a limited capacity whereas long term memory is considered unlimited. According to the theory, there are two types of cognitive constraints of working memory: *Intrinsic constraints* and *extraneous constraints* (Tardieu & Gyselinck 2003, 5).

Intrinsic cognitive load refers to the inherent difficulty or complexity of the presented material; how many elements it consists of and how they interact with each other (Mayer 2001, 50). The more elements there are and the more they are in interaction, the higher the intrinsic cognitive load becomes because the language learner has to attend to more elements simultaneously (Tardieu & Gyselinck 2003, 5). The intrinsic cognitive load decreases as the language learner gains more expertise; they become able to deal with configurations of interacting elements as single units (Paas & Kester 2006, 282).

The extraneous constraints refer to the way in which the material is presented; simultaneous modes that convey the same information as well as extraneous or irrelevant information increase the extraneous cognitive load (Mayer 2001, 50). By default, only one mode of media is required to present information. Thus, when two media present the same information, the other one is redundant and additional

processing which increases the cognitive load is required (Tardieu & Gyselinck 2003, 6).

The cognitive load is ineffective if it interferes with schema acquisition and effective if it fosters schema formation and automation (Paas & Kester 2006, 283-284). Working memory is constrained in that if the intrinsic and cognitive load together exceed its capacity, learning or understanding will not occur (Paas & Kester 2006, 283). This is due to *cognitive overload*, the exceeding of the capacity of working memory.

Coherence principle is also useful in understanding the constraints of working memory. Coherence principle has it that any extraneous information that does not help convey the message of the text intends to convey, hinders the understanding of the text. According to *Arousal Theory*, adding interesting pictures, music and sound effects makes the audiovisual text more enjoyable for the perceiver, thereby increasing the perceiver's emotional arousal and thus directing attention towards the audiovisual text (Mayer 2001, 119). Then again, these extraneous sources of information can distract the organization of the material in working memory and can lead the language learner to organize the material around an inappropriate theme (Mayer 2001, 124).

Another key concept is the *split-attention effect*. This effect takes place when the language learner has to divide their attention among and mentally integrate multiple sources of information (Tardieu & Gyselinck 2003, 8). Attention is a limited resource (Baltova 1994, 510) and dividing it will result in poor understanding of the presented material. Simultaneous presentation of verbal and visual information is beneficial when the material is combined into a unitary source; this allows learners to avoid dividing their attention between several sources of information, encourages them to pay attention only to relevant aspects of information and allows them to hold more information in each working memory store, reducing the processing load of working memory (Tardieu & Gyselinck 2003, 17-18).

The *contiguity effect* further accounts for the understanding of a multimedia text. This effect has to do with the presentation of the text. Simultaneous modalities explaining the same object of learning prompt learning more than two modalities presented separately in succession (Tardieu & Gyselinck 2003, 9). It is easier to draw connections between the two modalities when they are simultaneous. The contiguity effect is not to be confused with extraneous load caused by two modes conveying the

same information which hinders schema acquisition. The contiguity effect has to do with two modes that co-operate in conveying the information.

2.5 Nonverbal communication

In order to establish coherence and determine causal relations of the presented information, a language learner's comprehension depends on explicit information as well as on the recognition of the speaker's implicit and unstated but logically implied message (Dahl & Ludvigsen 2014, 814). Explicit information refers to speech that does not require special interpretation whereas implicit information can be something that has to be read between the lines or something conveyed via other means than spoken language. Nonverbal communication can consist of body movements, head nods, hand-arm gestures, facial expressions, eye gaze, posture, interpersonal distance and movement of lips (Sueyoshi & Hardison 2005, 661-662).

Non-verbal communication may have many communicative functions; it provides an alternative expression of the same speech act that can help capture attention, it provides message redundancy (i.e. presenting the same information more than once), it establishes a connection between the physical environment and speech and it may help resolving ambiguous messages and ease comprehension of complex semantic information (Dahl & Ludvigsen 2014, 814).

Even though nonverbal communication is helpful, it may not always be understood as implied; for example, cultural differences in the use of nonverbal communication such as gestures may result in wrong interpretations of the intended message (Baltova 1994, Sueyoshi & Hardison 2005). Therefore, some familiarity with the speaker's culture is required of the listener if they are to understand the nonverbal communication correctly, unless of course the speaker uses gestures typical to the listener's culture.

Gestures are quite possibly the most obvious form of nonverbal communication. Therefore, some basic theoretical concepts of gesture studies that are relevant in the light of this research will be explained. A definition of gesture is in place: gestures are often spontaneous movements of the hands and arms that are comprised while speaking in order to convey images that are isomorphic with what is being expressed verbally or to supplement ideas that are not otherwise expressed (Dahl

& Ludvigsen 2014, 813). However, gestures are not only intended for the listener, but they may be unconscious and help the speaker recall lexical items faster (Sueyoshi & Hardison 2005, 663).

Gestures can be classified to a multitude of different types. A few most common classes that appear in the texts used in this study will be identified here. *Iconic gestures* are gestures associated with meaning; they are used most often when a speaker is describing specific concrete things (Sueyoshi & Hardison 2005, 664). *Metaphoric gestures* are visual images, but they relate to more abstract ideas or concepts (Sueyoshi & Hardison 2005, 664). *Beat gestures* are associated with the rhythm of speech; they do not describe anything, but are used to help the speaker control the pace of speech (Sueyoshi & Hardison 2005, 663).

2.6 Previous studies on understanding audiovisual texts

In this section, the research methods and results of previous studies on the area of audiovisual listening comprehension will be discussed. First, the used research methods will be dealt with, then a look into the general research results will be taken after which the explanations given for the results by the authors will be discussed, then research results associated specifically with gestures will be analyzed and finally pedagogical implications of previous studies will be discussed.

A very common technique used in researches concerning the effects of multimodality on listening comprehension of EFL is to expose groups of learners to the same text but with a differing amount of modalities. One group may only hear the audio, another group also sees the video and a third group may have subtitles inserted to the video. By comparing the results of a comprehension test, conclusions can be drawn of the effect of each added modality. It should be borne in mind, however, that the added modalities are not the only factor that has an effect on the results, but it is also possible that there are differences between the language proficiency of the subject groups. Some of the studies referred to are case studies and some are researches of a much larger scale, but as we shall shortly see, the results follow clear tendencies. Therefore, I would suggest that the tendencies in these results show a more generalizable pattern of phenomena in comprehension of audiovisual texts.

A very clear tendency is that the test results of EFL learners are significantly better in a condition in which they both hear the audio and see the video when compared to the condition in which they only hear the audio of a text (Sueyoshi & Hardison 2005, Guichon & MacLornan 2008, Wagner 2010, Liang 2013). An exception was found in the study of Batty (2015), but this seemed to be due to the fact that Batty's test was designed so that it did not make effective use of the nonverbal channel which the researcher himself points out. Even so, Batty's test subjects performed slightly better in the audio and video condition than in the audio-only condition.

Although the general effect of video on listening comprehension performance is of much interest, the explanations that the researchers give for the gained test results and specific observations that the researchers have made are of at least as much importance. Guichon and MacLornan (2008) found out that if the visual information is not directly linked to the audial information, split-attention effect and cognitive overload are likely take place just as is theorized by Tardieu and Gyselinck (2003) and Paas and Kester (2006). Liang (2015) found out that visual cues benefit the activation of the viewer's background schemas in two ways: through the speaker's body-language and the environment shown in the video. Wagner (2010), too, found out that both the speaker's body-language and visual cues that do not relate to the speaker seemed to be beneficial for the test takers. Liang (2015) also noted that sometimes visual cues distracted the viewers, a finding made also by Baltova (1994).

Wagner (2010, 2) and Baltova (1994, 517) argue that because of its attractiveness to language learners, seeing the video helps them concentrate more than hearing an audio-only text. Baltova backs this argument up by the fact that none of her test subjects stated that they enjoyed listening to the texts in the audio-only condition whereas 70% of the subjects in the video condition enjoyed the texts (Baltova 1994, 518). Wagner's research results show that his test subjects performed better in the video with audio condition than in the audio-only condition even when it came to questions where visual help was not available (Wagner 2010). This could suggest that the subjects in the video and audio group were more proficient in the language, but this effect of group variation seems rather unlikely since Wagner's sample was relatively large; both groups consisted of approximately 100 randomly chosen participants. Therefore, it seems likely that the subjects in the video with audio group were able to concentrate on the text more intensely.

Dahl and Ludvigsen's (2014) research results show that EFL learners significantly benefit from seeing the speaker's gestures whereas native speakers do not. This is probably due to the fact that native speakers already understand the text nearly completely without seeing the gestures. Dahl and Ludvigsen's EFL learner subjects actually reached a level of comprehension comparable to that of the native speakers when they could see the speaker's gestures (Dahl & Ludvigsen 2014). Sueyoshi and Hardison (2005) made similar observations; lower proficiency EFL learners greatly benefited of seeing the speaker's gestures but for higher proficiency learners, the movement of lips seemed to be even more beneficial. They base this conclusion on the fact that higher proficiency learners gained best test results in a condition in which they only saw the speaker's face whereas the lower proficiency learners gained best test results in a condition in which they also saw the speaker's gestures (Sueyoshi & Hardison 2005).

Going beyond just the comprehension of audiovisual texts Secules et al. (1992), Weyers (1999) and Kon (2002) found out that the use of authentic video material in listening comprehension tasks prompted learning of a foreign language. Kon (2002) found out that his subjects were able to learn new vocabulary efficiently through the use of video as learning material. Weyers (1999) conducted a longitudinal study in which listening comprehension test results of two groups were compared to each other. The experimental group used authentic video material regularly in the classroom whereas the control group relied on more traditional methods. In the end, the experimental group's performance rose significantly more than that of the control group in listening comprehension. In the similar longitudinal study of Secules et al. (1992), the research results were matching to Weyers' study. However, Secules et al. noted that the skills the learners acquired were task specific (Secules et al. 1992, 486-487); unlike Kon's (2002) subjects, these learners did not benefit in learning new vocabulary or grammatical structures. The subjects specifically learned listening skills. Therefore, it can be said that using authentic video as learning material is beneficial for learning listening comprehension but there is controversy on its usefulness for teaching other skills.

3 Understanding accents

The cognitivist theory of information processing has it that all new information is screened through, or checked against existing information structures. When it comes to processing speech in particular, there are three categories of factors that have to be taken into account: 1) *linguistic or input factors* such as phonetic-phonological, lexical and syntactic features of the speech signal, 2) *sociolinguistic or interpersonal factors* such as the listener's attitudes and expectations, 3) *psycholinguistic or knowledge systems* such as world knowledge, linguistic knowledge and shared knowledge (Pihko 1997, 20). This section will mostly deal with the first one of these categories with the role of accent in mind. Furthermore, the role of linguistic knowledge, which is a part of the third category, will also be covered. The second category is a rather large topic and one of the main aspects of this research. Therefore, it will be explained in detail in section 4. When it comes to understanding speech, world knowledge functions in a similar fashion as it does when understanding an audiovisual text: top-down and bottom-up processing are used to connect previous knowledge with the presented information. Therefore, I will not deal with the cognitive theory behind world knowledge in length as it was quite thoroughly explained in the previous section.

A definition of the word 'accent' as to what it stands for in the context of this study is in place. An accent is a way of speaking that every speaker has; it consists of the use of particular consonant and vowel sounds as well as rhythmic, intonational and prosodic features; it includes the interrelationships between these features and sounds and the rules that control their use (Wells 1982, 1). The difference between accent and dialect is that accents only include features of pronunciation whereas dialects include specific sets of syntax, morphology, lexicon and pronunciation (Wells 1982, 3).

3.1 Prosodic features

Each accent and each speaker has a different set of prosodic features that influence the intelligibility of speech for the listener. These features include speech rate, pausing, tone, stress and rhythm. Some features may make words unrecognizable for the listener whereas some may cause a delay in recognition. It is important to bear

in mind, however, that no accent is objectively more intelligible than any other accent; as will be explained later on in this section, the listener's experience in and accustomedness to hearing a specific accent will make the accent in question easier for the listener to understand whereas an accent with which the listener is not familiar with can be more difficult to cope with.

Speech rate is sometimes considered a part of prosody, although sometimes it is handled as a separate feature. In this thesis, I will discuss it as a part of prosody. Speech rate is one of the most important linguistic factors that affects intelligibility because it can have an influence on other factors either directly or indirectly. Unsurprisingly, fast speech rate can hinder understanding whereas a slow speech rate is most often easily intelligible. This is because fast speech rate leaves little time for processing the input (Pihko 1997, 45-46). Non-native speakers tend to speak more slowly than native speakers (Pihko 1994, 100-101) which is the reason why fast speech rate is a factor that can make an accent less intelligible especially when it comes to native accents.

Pausing is in close connection to speech rate. Actually, speech rate is considered to consist of articulation rate and the number and length of pauses. Pausing can either facilitate or hinder intelligibility; when pauses are used in a syntactically logical manner as junctures, even fast speech becomes more intelligible but when pauses are illogical, even the intelligibility of slow speech can decrease (Pihko 1997, 46). Such a case could be for example if the speaker hesitates or stutters. Illogical pauses can also be caused by the lack of language proficiency on behalf of the speaker; they may need to spend more time to think of what to say and sometimes may have to pause to correct themselves or to rephrase a sentence. Thus, pausing can cause difficulties for the listener especially if the speaker is a non-native speaker of English, although this does not by any means mean that pausing could not cause difficulties when listening to native accents, as well.

Intonation, tone, stress and rhythm are also significant factors of intelligibility. A knowledgeable EFL learner may find these prosodic features useful but for a beginner they may be confusing as they can differ significantly from those of the learner's mother tongue (Pihko 1997, 46). Reductions are another prosodic input factor that also has significance for intelligibility. They are simplifications that often take place in informal speech. Examples are weakening of vowels and consonants, contractions, assimilations, elisions and liaisons (Pihko 1997, 46). Reductions may be

a feature of both native and non-native accents and often make the speech signal hard to understand for a listener who is not accustomed to them (Pihko 1997, 46).

3.2 Knowledge systems

The relation between the linguistic knowledge system of the listener and lexical features of the speech signal are best explained by the theory of Floccia et al. (2006). Human brain needs to match the incoming language input to a myriad of acoustic forms stored in the lexicon in order to understand speech (Floccia et al. 2006, 1276). In other words, an utterance can be understood completely if the listener's stored lexicon includes the lexical features, i.e. the words, used in the speech signal. When a person hears an accent which they are not familiar with, complex normalization mechanisms are needed to match the input to stored acoustic forms (Floccia et al. 2006, 1277). The further the input is from the stored acoustic forms, the more complex the normalization mechanisms need to be (Floccia et al. 2006, 1277). There are two distinct phases to this normalization: an initial period in which comprehension is disrupted, and an adaptation phase which leads to total or partial recovery of comprehension (Floccia et al. 2006, 1277).

Clarke and Garret (2004) found out that the initial period most often takes around one minute of exposure to be passed. After that the reduction in reaction times starts to diminish and the adaptation phase can be accessed. The delay in the recognition of words caused by an unfamiliar accent hinders understanding because when the recognition of individual words takes too much time, there may not be enough time to decode the whole utterance.

Another essential knowledge system is *shared knowledge*. It refers to the knowledge of the topic of speech, the type of speech situation, knowledge about the speaker and their background, speech style and voice quality (Pihko 1997, 19). All of these factors influence intelligibility as they can be used for top-down processing.

3.3 Coping with accents

Now that it has been established that accented speech causes difficulties in intelligibility for language learners by making familiar words unrecognizable or at

least by producing a delay in the recognition of words, it is time to explain how these difficulties can be overcome. In the light of their research results, the researchers on the field of understanding accented speech are unanimous in the conclusion that experience in hearing an accent is crucial in making the heard accent more intelligible (Smith & Bisazza 1982, Gass & Varonis 1984, Pihko 1994, Pihko 1997, Bradlow & Bent 2003a, Bradlow & Bent 2008, Kennedy & Trofimovich 2008, Adank et al. 2009).

The research results of Bradlow and Bent (2003a & 2008) show that being exposed to a single speaker of a non-native accent helps the listener develop a better understanding of this specific speaker whereas exposure to multiple speakers with the same language background develops a speaker-independent adaptation to the accent. To be exact, training with multiple speakers of the same accent enabled the listener to understand an unfamiliar speaker of the same accent as effectively as training with the speaker in question would have (Bradlow & Bent 2003a). Furthermore, the research results showed that the lower the baseline understanding of the accent the listener had, the more exposure they needed in order to adapt to the accent (Bradlow & Bent 2008).

Gass and Varonis (1984) as well as Kennedy and Trofimovich (2008) highlight the role of background knowledge as a helpful tool for understanding accented speech. Indeed, background knowledge may relieve the burden of working memory by patching up ununderstood gaps in the speech signal. However, background knowledge is not something that could be thoroughly practiced as a part of language learning and thus gathering general knowledge seems an inadequate means of enhancing the understanding of accents.

3.4 The interlanguage speech intelligibility benefit

A phenomenon that helps explain some of the research results gained in the area of accent intelligibility is the *interlanguage speech intelligibility benefit* (ISIB). ISIB can be divided into two types: *matched* ISIB and *mismatched* ISIB. Matched ISIB means that foreign language speakers of the same language background find the speech of each other easy to understand. One explanation for this is that non-native speech production and perception are often both systematically linked to native language sound structure (Bradlow & Bent 2003b, 1607). Of course, the speakers of the same language background are most probably also very accustomed to hearing the accent of

their language background which gives them a benefit in understanding the language because of the familiarity of the accent. As Pihko puts it: “For EFL learners, presumably, the most intelligible speaker is often their own English teacher,” (1994, 40).

Mismatched ISIB on the other hand means the benefit foreign language speakers get when listening to the speech of foreign language speakers of a different language background in comparison to native speakers listening to foreign accents. This is explained by the lack of prosodic reduction phenomena that often appears in the speech of non-native speakers (Bradlow & Bent 2003b, 1602). Pihko (1994, 100-101) states that non-native speech tends to be characterized by features that have the potential to make non-native speech more easily intelligible to non-native listeners: fewer assimilations, word-linkings, elliptic forms, elisions and vowel and consonant weakenings and word boundaries tend to be clearer than in native speech.

The research results of Bradlow and Bent (2003b) support the existence of both matched and mismatched ISIB phenomena, but they arouse controversy among researchers and quite justifiably so; it is apparent that certain prerequisites have to be fulfilled so that ISIB can take place. For example, the foreign language speaker has to pronounce the consonants and vowels quite clearly in order to other non-native speakers to distinguish them and be able to have any use of mismatched ISIB. If the speaker does not pronounce sounds clearly but in a way unexpected for the listener, cognitive normalization mechanisms need to be used in order to match the input to the stored lexicon and a delay in processing will take place. Even though non-native speech has the potential to be more easily understandable for non-native listeners, non-natives' lack of proficiency may restrict their ability to correctly interpret speech that deviates strongly from the norm that they are used to (Pihko 1994, 100-101).

3.5 Previous research on understanding accents

Results in studies that have tried to determine which accents are easiest to understand for EFL or ESL learners show that the easiest varieties are either the learner's own variety (Wilcox 1978, Moinzadeh et al. 2012) or the standard variety that has been used most at school (Smith & Bisazza 1982, Pihko 1997). These results can be explained by the effect of familiarity of accent especially in the cases in which standard

varieties have been found to be the easiest accents to understand, and by the ISIB as well as familiarity in those cases in which the subjects' own variety was the most intelligible. Going beyond simply trying to determine which accents are most difficult to understand for EFL learners, Pihko (1997) found out that the less proficient the learners were in English, the harder time they had coping with unfamiliar accents. However, even high proficiency learners showed significant difficulty with unfamiliar accents which suggests that the flexible ability to understand different accents of spoken English does not develop automatically in connection with general EFL proficiency (Pihko 1997). It seems quite clear that students need instruction and experience in listening to varieties of English that deviate from the standard varieties used at school.

Some of the researches referred to in this section focus on how native speakers understand accented speech, but as Major et al. (2005, 58) concluded, it is likely that native speakers and non-native speakers are affected in the same way by accented speech. Therefore, these researches can be considered relevant in explaining non-native speakers' processes of understanding accented speech. Furthermore, as language learners' understanding of target language accents seems to be a surprisingly little studied subject, this research relies on all of the research information available.

The amount of relevant previous research in the light of this study is further narrowed down because some researches on this area have studied perceived comprehensibility rather than actual intelligibility. Derwing and Munro (1997) studied the relation between the listeners' idea of their understanding of accented speech and their actual understanding. They found out that even fully intelligible accented speech may require additional effort or processing time of the listeners, which often leads them to rate accented speech as more difficult to understand than it actually is (Derwing & Munro 1997, 12). Thus, the results of studies that do not take the difference between intelligibility and comprehensibility into account cannot be considered valid sources of information when focusing on intelligibility.

4 Attitudes towards accents

Listeners' subjective evaluations may be an intelligibility factor as critical as the phonetic-phonological characteristics of different varieties of English (Pihko 1997, Major et al. 2005). This section explores the reasons behind certain types of attitudes as well as the effects that attitudes may have on intelligibility. As we shall shortly see, sociolinguistic and interpersonal factors such as listener's attitudes and expectations play a major role in understanding accents. Since this section will discuss attitude from multiple points of view, a definition is in place. Allport defines attitude as follows: "An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related." (cited in Gardner 1982, 132)

4.1 Why does attitude matter?

The reason why attitude is so important in understanding accents is that listeners with negative attitudes towards a language variety may either reject the speech or at least tend to make less of an effort to understand the speech they deem difficult to understand (Pihko 1997, 50). This can be seen in the research results of Eisenstein & Verdi (1985): African American Vernacular English was considered difficult to understand by the ESL test subjects who were very familiar with the variety but had a negative attitude towards it. Because of their familiarity with the accent, they should have been fully capable of understanding the speech signal, yet they reported that they did not understand it.

Derwing and Munro (1997) found out that some accented but fully intelligible utterances may require additional effort or processing time, which leads listeners to rate them as difficult to understand. Furthermore, the research results of Lopez-Soto and Barrera-Pardo (2007) show that language learners' attitudes are more negative towards varieties they find difficult to understand. In other words, attitudes easily become negative towards accents that are considered difficult to understand because they require more effort on the listener's behalf. This may happen even if the heard accents were actually perfectly intelligible. Going beyond just the effect on

understanding accents, Bradac and Giles point out that students' language attitudes may either facilitate or inhibit their learning of a second language (Bradac & Giles 1991, 2).

Pihko identifies a phenomenon called 'listener irritation' which refers to the affective reactions of annoyance that deviant and non-standard speech arouses in the recipient (Pihko 1994, 21). According to her, listener irritation can lead to unwillingness to co-operate with the speaker to such degree that the listener does not even want to understand what they are saying (Pihko 1997, 125). On the other hand, Pihko's research results show that positive attitude towards an accent prompts better understanding (Pihko 1997). A positive attitude towards a certain variety of English prevents listener irritation thus ensuring that the listener does not give up on trying to understand the speaker.

4.2 The formation of attitudes

When studying language learners' attitudes towards accents, it is essential to understand how these attitudes are formed. Attitude formation towards accents will be explained following Bradac and Giles' (1991) rather comprehensive theory quite closely but also associating some ideas of other researchers who have studied the same area. EFL learners are likely to evaluate accents of English along the same way they would evaluate accents of their native language (Bradac & Giles 1991, 9). But how are attitudes towards accents formed in general then?

Two major theories have been proposed: the *inherent value hypothesis* suggests that some language forms are inherently better than other forms perhaps because of aesthetic characteristics (Bradac & Giles 1991, 6). However, this theory has never gained any serious support since it is very hard to believe that every speaker of a language shares the same idea on aesthetics. A more popular view is the *associationist explanation* which assumes that the features associated with language are not natural and inevitable but instead arbitrary and mutable (Bradac & Giles 1991, 6).

Bradac and Giles (1991, 2) name two primary evaluative dimensions of language attitudes which are *status* and *solidarity*. Status refers to characteristics such as intellect, sophistication and education etc. whereas solidarity refers to personal characteristics such as friendliness, trustworthiness and humorousness etc. Each of

these dimensions has two poles, one of which focuses on the individual speaker and the other on the speaker's group affiliations.

These poles can be made salient by a variety of factors in a given situation. For example, a speaker whose accent has a prestigious status in society may be considered high in the status dimension but low on the solidarity dimension by a speaker of a lower prestige variety. This speaker of a lower prestige variety is likely to consider another speaker of the same lower prestige variety high in the solidarity dimension but low on the status dimension because the variety they speak lacks prestige. On the other hand, two speakers of the same prestigious variety are likely to consider each other high in both status and solidarity dimensions because their accent has the required prestige and they identify with each other because of their similar accents.

People's evaluations of others in the solidarity dimension rely a lot on how much the listener identifies with the speaker. The status dimension, however, has to do with social prestige which an accent can attain by being associated with a socially powerful group who speak in this manner (Ladegaard 1998, 267).

Social in-groups and out-groups are essential factors in the formation of attitudes towards accents (Bradac & Giles 1991). In-group means a social group with which a person identifies and out-group is a social group with which the person does not identify. Bradac and Giles (1991, 5) define three basic models for attitude formation: 1) Speaker's language use leads to an inference about their group membership and stereotypical ideas of how people are in this specific group are associated with the speaker. 2) Speaker's language use leads directly to inferences about their personality. 3) Speaker's language use leads to an inference about their group membership, but the listener does not associate this group with any personality-related characteristics.

The first model is likely to take place for example when the listener has a strong in-group identity and perhaps pride and a belief that their way of speaking has institutional support, but the speaker belongs to an out-group (Bradac & Giles 1991, 4-5). An example of the first model could be a situation that takes place in Scotland and in which the speaker speaks with Received Pronunciation accent but the listener's accent is Mainstream Scottish English. The Scottish listener's own accent has social prestige in Scotland which makes the speaker a member of a somewhat rival out-group since Received Pronunciation is a prestigious accent within the same country but in another area. The listener associates the speaker's accent with the English upper class

thus making an inference about the speaker's group membership because of their language use. The listener considers British upper class people educated and competent but perhaps self-satisfied. Because the listener thinks the speaker belongs to the British upper class, they associate these personal characteristics with the speaker. Bradac and Giles' theory is supported by the research results of Bresnahan et al. (2002): they found out that people with a strong ethnic identity tend to favor their own variety of English over other varieties.

The second model may take place for example when the hearer does not have a strong in-group identity and they identify with the speaker's out-group (Bradac & Giles 1991, 5). An example of the second model could be a situation that takes place in London and in which the listener is a speaker of Midwest American accent and the speaker is a speaker of Welsh accent. The listener does not have a strong in-group identity because their in-group of Midwest American accent speakers in London is rather small in number. They identify with the speaker of Welsh accent because they too belong to a group who do not speak the same way as the locals. The listener may associate positive characteristics of the solidarity dimension such as friendliness or relaxedness with the speaker. The research results of Bresnahan et al. (2002) support Bradac and Giles' theory regarding to the second model, as well: in their research, people exhibiting weaker ethnic identity were found to be more receptive to accents other than their own.

The third model can take place for example if both the speaker and listener are members of the same in-group. An example of the third model could be a situation in which the listener and speaker are speakers of the same accent in a place where their accent is the dominant and prestigious accent. Because they are both members of the same in-group and encountering another member of this in-group is an everyday situation, the only assumption the listener makes about the speaker because of their accent is that the speaker is a local person.

As Bradac and Giles (1991, 5) state and as can be seen in the examples above, language evaluation plays a lot upon stereotypes and learned models. Personal characteristics associated with the people who belong to a certain group are of course stereotypical and stereotypes have to be either learned or formed in accordance to experience. This applies to ideas associated with speech styles, as well. For example, assuming that slow speech rate has to do with lack of intelligence is a learned model rather than an inherent characteristic of slow speech.

The fact that learned models and stereotypes are subject to change explains the arbitrariness and mutability of the associationist explanation of attitude formation towards accents. The situations in which Bradac and Giles' models take place are not static either: as will be explained in the following paragraphs, there are more factors that affect attitude formation and can make it more unpredictable.

Evaluative reactions to language are strongly affected by the context of communication (Bradac & Giles 1991, Dalton-Puffer et al. 1997). For example, rapid speech rate may be associated with intellectual competence and slow speech rate might be associated with the lack of intelligence. However, if the listener assumes that the speaker speaks slowly intentionally, they may appreciate the speaker's rhetorical sensitivity and therefore consider the speaker intelligent regardless of the slow speech rate (Bradac & Giles 1991, 3-4). An example of such case could be a situation in which the speaker talks to an audience consisting of children about a complicated topic – for example a teacher speaking to students. The listener assumes that the speaker speaks slowly intentionally so that his or students can keep up.

This does not directly relate to forming an attitude towards an accent but rather to forming an attitude towards the speaker. However, the attitude towards the speaker may play a major role when forming an attitude towards their accent. As Bradac and Giles point out, in some contexts, the hearer might evaluate the speaker based on assumed personal characteristics whereas in some other contexts the group characteristics may be more important for the listener (Bradac & Giles, 5). People may form an attitude towards a person on account of the person's accent but they may also form an attitude towards an accent on account of the people who speak this accent.

According to Bradac and Giles (1991), convergence and divergence are important, as well. Attempting to accommodate one's speech to match the listener's accent is often seen as a friendly act whereas diverging intentionally by moving away from the other's conversational behavior is considered socially distancing or even rude (Bradac & Giles, 3). Thus the solidarity dimension of the speaker is evaluated more favorably if they try to converge to the listener and more negatively if the speaker tries to diverge from the speaker. This, of course, can only happen in a communication situation in which the participants are in interaction.

4.3 Accent attitudes and EFL learners

Although being quite comprehensive, Bradac and Giles' theory can be extended on the basis of research results of other researchers. EFL learners' attitudes towards different accents of English reflect the learned attitudinal models of their native language. However, this does not fully explain the attitude forming behavior of EFL learners. EFL learners not only reflect the learned attitudes relating to their native language but they also reflect the ideas of the target language that they learn from their teacher (Rashid 2011, 74).

The effect of schooling extends even further: according to the results of several attitudinal researches on EFL learners, they tend to evaluate the accent used as pronunciation model on English lessons most favorably in both status and solidarity dimensions (Matsuura et al. 1994, Ladegaard 1998, Chien 2014, Rindal 2014, Buckingham 2015). Moreover, EFL learners tend to evaluate non-native varieties less favorably than standard native varieties especially when it comes to the status dimension, but the gap between non-native and native accents tends to be narrower when it comes to the solidarity dimension (Matsuura et al. 1994, Chiba et al. 1995, Chan 2013). These tendencies can be explained by another tendency: EFL learners tend to evaluate accents that they are familiar with more favorably than those they have not heard before (Matsuura et al. 1994, Chiba et al. 1995, Dalton-Puffer et al. 1997, Chien 2014). It is quite likely that EFL learners are not familiar with all the non-native accents they had to evaluate in the researches mentioned. Therefore, they evaluated more favorably the standard varieties which they certainly were familiar with.

Pihko (1997) and Hakala (2007) found out that EFL learners tend to evaluate negatively the accent that is typical to speakers with the same native language as they do. This has to do with the effect of schooling on attitudes towards accents; EFL learners' own accent is considered wrong because it is the accent they have tried to rid themselves of if they are encouraged to strive for a native-like accent at school.

Quite interestingly, the speaker's proficiency in English does not seem to have any significant effect on the speaker evaluation (Matsuura et al. 1994). Neither does the heaviness of accent (Cargile and Giles 1997). However, this research result can be considered to extend to EFL learners only cautiously since Cargile and Giles' subjects were native speakers.

Giles (1970) found out that people tend to evaluate non-native accents in a similar fashion as they evaluate regional native accents. However, I suggest caution when applying this result to EFL learners since the subjects in this research too were native speakers. On the other hand, there is considerable reason to believe that the behavior patterns that apply to native speakers apply to EFL learners, as well; Bradac and Giles (1991, 9) theorized that EFL learners are likely to evaluate accents of English in a similar fashion as they evaluate accents of their native language.

4.4 Techniques for testing accent attitudes

When designing a test to research attitudes regarding accents, it is essential to know what kind of different approaches to these kind of tests there are. Language attitudes are divided in two major evaluative dimensions: status and solidarity. However, this has not always been the case, but it is rather a product of longer development. In an early accent attitude research Giles (1970, 213) proposed that the evaluation of an accent includes three major categories: 1) aesthetic: the pleasantness-unpleasantness of an accent to the hearer, 2) communicative: the “comfort” of listening or intelligibility of an accent to the listener, 3) status: the prestige the accent holds in the listener’s view. Later on the communicative category has been most often combined with the aesthetic category because it has been shown that subjective comprehensibility does not reflect actual intelligibility. Thus, the intelligibility category has been deemed redundant.

Ryan et al. concluded that the two major dimensions along which views about accents vary can be termed social status and group solidarity (Ryan et al. 1982, 7-8). In their 1991 study Bradac and Giles agreed with this view stating that the two primary evaluative dimensions of language attitudes are status and solidarity (Bradac & Giles 1991, 2). Most modern researches that deal with attitudes toward accents categorize different aspects of attitudes along these two categories although the names of the categories may vary slightly.

There are two major techniques with which the material for an accent attitude test can be produced. According to Hakala (2007, 9), the most popular method has been the Matched Guise Technique (MGT). In this technique the same speaker produces multiple accents that are used as examples that are played to the subject

group. The subjects are often led to believe that each accent is spoken by a different speaker. The positive aspect of MGT is that it eliminates possible attitudes that have to do with the speaker's personality rather than their accent. The negative aspects are that the produced accents are not authentic nor is the communication situation, because the recording of the accents has to be thoroughly planned in advance. Furthermore, it may be difficult to find a speaker who can produce several credible accents.

Another technique that was developed as the idea of the MGT was elaborated is the Verbal Guise Technique (VGT). In this technique there are two major differences when compared to MGT: the use of authentic material is possible and several speakers can be used (Hakala 2007, 12). The advantages of VGT are obvious: when several people can be used to produce the desired accents, real speakers of these accents can be chosen which results in authenticity. It may also be easier to find speakers of various accents than one speaker who could produce multiple accents. The communication situation, or the recording situation, does not have to be planned in advance because one person does not produce all accents. This too enables authenticity. Nowadays with the use of Internet, an almost infinite amount of authentic material can be found. The only disadvantage is that VGT does not eliminate the attitudes regarding the speaker's personality. The role of the speaker's gender has been speculated; some researchers have deemed possible that attitudes towards speakers of different sex may be significantly different. Rindal (2014) supported this view in the light of her research results. On the other hand, Rashid's (2011) research results showed that the gender variable had no significant influence on subjects' preferences. It remains unclear which view is correct and it may well be that they both are; the test-subjects' attitudes towards genders may vary depending on for example their cultural backgrounds.

The subject group's answers can be gathered in various ways; open-ended questions, interviews, bi-polar scales and Likert scales have been used. However, attitude questionnaires can be divided into two major classes: attitudes towards accents can be elicited either by the use of direct questions about the language variety or by inferring attitudes indirectly from evaluations of the speaker's personality or attributes (Ryan 1982, 7-8).

Hakala (2007, 26) states that *indirect assessment* is nowadays the most often chosen method because it is thought to reveal spontaneous and thus presumably more genuine attitudes than *direct assessment*. Direct assessment has been previously a widely used method but later on its validity has been questioned because it has been

thought to have the disadvantage of possibly producing socially desirable answers (Hakala 2007, 26). However, Groelanders and Van Hout's (2010) research results proved that there is no difference in the results of direct and indirect assessment methods. They both seem to elicit subconscious attitudes rather than conscious and socially desirable ones.

5 Material and methods

In this section, the research design and the reasoning behind the choices made when designing the test will be discussed. This section starts with the discussion about the choice of texts for the purposes of this research. Then the choice of intelligibility testing method will be explained after which the structure of the attitude questionnaire will be opened up. Finally, the actual testing situation and sample are discussed.

5.1 Audiovisual texts

Four videos from the social media video site YouTube.com were retrieved for the research. The prerequisites for the videos included that there is only one speaker, the video must be from four and half to five minutes in length, there can be no background sounds or noise that could not be removed by editing and there can be no extraneous visual material that can either facilitate or hinder understanding the text. Furthermore, the speaker had to be visible from at least chest up in order for their hand gestures to be visible. All of the chosen videos were storytelling videos in which a single speaker told a story to the camera or to an interviewer who is not visible.

The texts used in this study make use of the following audial modes: sounds made by the speakers such as speech and other sounds produced by speech organs, tone and volume of speaking voice and sounds caused by the speaker's body-language such as clapping. The visual modes in this study consist of the speaker's body-language such as hand-arm gestures and facial expressions and a static background which bears no significance to the text. Thus, the audiovisual texts used in this study focus solely on the speakers. Therefore, their intrinsic cognitive load can be considered low as the test-takers do not have to focus on many simultaneous elements.

The accents on the videos are British Received Pronunciation (RP), Irish County Kerry accent, Zimbabwean native Shona speaker accent and Japanese accent. Therefore, there is one speaker for each of the following categories of English speakers: standard native English speaker, regional native English speaker, English as a second language (ESL) speaker and English as a foreign language (EFL) speaker. As English is an official language in Zimbabwe, the Zimbabwean speaker is the ESL

speaker. The accents of the speakers were not determined only by listening to the characteristics of their speech but also by finding information on the speakers and their backgrounds. As the speakers are more or less professional storytellers, it was relatively easy to find out the necessary information about them.

Since there are multiple speakers with authentic accents rather than one speaker who musters several accents, the test makes use of the Verbal Guise Technique (VGT). The Matched Guise Technique (MGT) is often considered advantageous in that it eliminates the factors on understanding and evaluating accents that the use of several speakers imposes. However, it is practically impossible to find a speaker who could produce several convincing accents when it comes to non-native varieties (Dalton-Puffer et al. 1997, 118). In addition, the obvious advantage of VGT is that authentic material can be used. Besides, the point of MGT tests is that the informants are not aware that the same speaker produces all of the texts. Since the informants can see the speaker in this test, it would be impossible to convince them that there are several speakers instead of one.

Storytelling was chosen as the topic so that no technical vocabulary would be used in the texts and the challenge in understanding would rely mostly on how the informants were able to cope with the accents. Furthermore, the vocabulary needed to be easy so that none of the texts would be inherently more difficult than others since it cannot be known for sure what kind of vocabulary the informants are familiar with. Another advantage with stories is that they do not have to be true. Therefore, the informants cannot solely rely on their world knowledge and competence in the test type, but their understanding of the text is indeed tested.

Storytelling also often incorporates vivid use of gestures and facial expressions. This physical body-language has the potential to make the visual mode useful for the informants as it gives the informants also visual cues of what is being said. The gestures on the videos include especially iconic, metaphoric and beat gestures. The different types of gestures were identified in section 2.5 of this study.

Another advantage of storytelling videos is their authenticity; the speaker knows the story by heart and tells it in their own words rather than reads it from a paper. Furthermore, the focus is on a single speaker. There are other possibilities for videos with a single speaker such as lecture tapes, news reading or podcasting, but there are some issues with these types of videos. News reading is most often not spontaneous since the reader indeed reads a written text out loud, lectures may

incorporate difficult vocabulary and podcasts do not necessarily include gestures in the same extent as storytelling.

It turned out to be impossible to find videos that would completely fulfill all of the prerequisites set for the texts. In the beginning of two of the videos there was a small text that stated the name of the speaker for a short duration. Black boxes that covered these texts were added to the videos, so that the informants could focus solely on watching the speaker while listening to the story. These boxes do not cover the speaker thus ensuring that no gestures are left unseen. As there is no other extraneous visual material in the videos and the written texts were covered with black boxes, the extraneous cognitive load of the videos can be considered minimal.

The video with the Japanese speaker was originally eight and half minutes long but quite conveniently, there was a very fitting break around four and half minutes to the story where a sort of second stage of the story began. The video was cut to include only the first stage of the story which serves the purposes of this study very well.

Since the videos have been filmed with different equipment, there is some variation in the sound and video quality. However, the soundtracks of the videos were edited with the sound editing software Audacity so that they correspond as much as possible. This was because background noise, for example, has been found to produce a delay in the recognition of words (Adank et al. 2009, 524). The sound quality of the videos with slightly poorer quality was enhanced by removing background hum and by amplifying the volume of speech. The videos were then all saved with the same resolution in order to ensure similar precision of picture. With these enhancements the possible effects the test material can have on the test results were minimized.

5.2 Multiple choice test

Multiple choice (MC) test with four answer options was chosen as the method to measure the participants' understanding of the text for various reasons, some of which relate to the positive aspects of MC tests and some to the downsides of other testing methods. First of all, Finnish EFL students are definitely familiar with this type of test because MC tests are very often used in the classroom and examinations. Therefore, it was considered unlikely that any problems relating to the understanding of the task would be encountered. MC test is also the most objective test method because if

designed properly, there is only one answer that is absolutely right and the rest of the answer options are simply wrong. There are two downsides to MC tests. The first and the more obvious one is the possibility to guess the correct answer. The other one is that they do not show the reasoning behind the choice of answer option (Cheng 2004, 545).

The amount of answer options was restricted to four for a few reasons. If there were three answer options, it would be too easy to simply guess the right answer. Having more answer options also makes the strategy in which the test-taker singles out the right answer more difficult. This strategy is problematic because the test-taker could choose the correct answer for an incorrect reason (Yi'ian 1998, 30). Thus, the easier it is to guess or single out the right answer the less the test can be trusted to actually test understanding. However, if there were five or more answer options, it would take the participants a long time to read through all of them and they would have to keep all of them in mind while listening to the text which would burden the working memory more heavily. This is undesirable since the purpose of the test is to measure understanding instead of the capacity of working memory. It has been suspected that in MC tests, the test-takers could find cues for the correct answer before hearing the text by reading the answer options (Cheng 2004, 550) but Yanagawa and Green's (2008) research results proved that this is not the case. Instead, showing the questions beforehand can be beneficial for the test-takers because they know what kind of information to look for in the text (Yi'ian 1998, 38).

The test itself was composed in Finnish. This ensured that the test-takers certainly understood all of the questions and answer options which is essential since misinterpretation of the answer options may facilitate the choice of wrong answers (Yi'ian 1998, 38). Cheng states that test-takers are able to demonstrate their comprehension more fully when they are tested in their native language because they can summarize and process the heard text in their own language (Cheng 2004, 547). When test-takers do not fully understand the heard text, they may often use a strategy of matching the answer options with items in the text (Yanagawa & Green 2008, 110). They can select an answer option because there seems to be a relation between the words that appear in the answer option and the text; the answer option may have a word or words that the test-takers heard in the text, or there could be words that are similar in sound or meaning in the text and the answer option, or the words in the answer option and the text may simply belong to the same word group (Yanagawa &

Green 2008, 110). When the text and answer options are in different languages, this strategy becomes a less apparent option for the test-takers.

The questions were designed so that they can be answered on the basis of the audio alone, but the body-language of the speaker may give additional cues for the test-takers. This was relatively easy since there is no visual material other than the speaker, meaning that there are no possible extraneous visual cues that would put the audio-only group in an unfair disadvantage. The answer options, on the other hand, were designed so that there indeed was only one correct answer option. The wrong answer options either included something that was not stated in the story or something that was stated in the story but did not answer the question. Many of the wrong answer options made use of world knowledge that could be thought to be relevant to the question but actually did not help answer it correctly. This was done so that the research results would show when the test-takers had to rely on top-down processing and when they could understand the text so well that only bottom-up processing was required in order to answer the questions.

Each text was divided into five short stretches of roughly equal length. Two questions were asked of each stretch and pauses were taken between the stretches so that they all would burden the test-takers' working memory to the same degree. The answers to the questions appeared at random locations in each of the corresponding text stretches but so that the answer to the first question was always before the answer to the second question. This was done because the test would become easier if the answers would always appear at the beginning and end of the text stretch (Yanagawa & Green 2008, 113). In this case, the test-takers could be able to predict when the correct answers would appear and they would not have to understand the whole text in order to answer the questions correctly.

There are alternative ways of testing listening comprehension but all of them have their downsides. Finnish EFL learners may not be very familiar with dictation and partial dictation tests. This could cause confusion among the test-takers which could affect the test results. Furthermore, these test types are not very suitable for audiovisual material because the test-takers cannot write and look at the video at the same time. Dictation tests also require that the test-takers produce language, which is not meaningful for this study. Dictation tests are also hard to evaluate objectively because the test-takers may produce a multitude of different kinds of errors all of which cannot be considered equally severe. Determining the boundaries between minor and

major errors can be very difficult and it adds the effect of the researcher's opinions on the test results. Cloze test and multiple choice cloze tests are out of the question for a similar reason as dictation tests; the test-takers cannot read the gapped written text at the same time as watching the video.

Like dictation tests, open-ended questions could possibly elicit some qualitative data and guessing would become more difficult, but they are hard to evaluate completely objectively, as well (Cheng 2004, 546). Since the test-takers could come up with answers that are partially correct but partially wrong, a more complicated evaluation method would have to be designed. Using open-ended is thus a method which inevitably would include some amount of subjectivity as the researcher would have to draw lines between ambiguous answers. Open-ended questions also burden the working memory more than MC tests because more information has to be held in it in order to be able to write down the correct answer (Cheng 2004, 550).

5.3 Speaker evaluation questionnaire

A questionnaire was developed to measure the test-takers attitudes toward the accents of the speakers in the MC test. This questionnaire enabled the participants to evaluate the speakers on three status-related aspects as well as three solidarity-related aspects. The status related aspects were intelligence, sophistication and education and the solidarity-related aspects were friendliness, humorousness and honesty. These specific characteristics were chosen for they were found out to be the most often evaluated characteristics in researches focusing on attitudes toward accents. Therefore, comparisons between this study and previous ones can be made.

These six aspects were evaluated on a bi-polar scale with opposite adjectives on each end and seven numbers in between. A number of possible values as high as seven was chosen because the sample size was small. Thus, it was ensured that variance would appear. The number was uneven so that a neutral value could also be given if the accent did not arouse neither positive nor negative feelings in the test-takers. Open-ended attitude questions were not included because comparing attitudes described freely could prove difficult. When using a clear scale, the attitudes could be easily compared to each other and mean scores could be calculated. The aim was once again objectivity. Furthermore, the research results of Grondelaers and Hout (2010)

have it that the use of speaker evaluation scales does not affect the conscious nor subconscious attitudes of the test-takers. In addition, they found out that free response attitude tests elicit same kinds of adjectives that are often used in speaker evaluation tests which make use of bi-polar and Likert scales (Grondelaers & Hout 2010).

Indirect assessment is nowadays the most often used method of attitude evaluation tests (Hakala 2007, 26). In indirect assessment, the assessors are asked to evaluate the speaker's characteristics on the basis of a heard audio sample of the speaker without mentioning that the results are interpreted with accent in mind. Indirect assessment has been considered to be able to elicit underlying attitudes rather than conscious ones, whereas direct assessment has been thought to include the risk of producing socially desirable answers (Hakala 2007, 26). However, Grondelaers and Hout (2010) found out that this is not necessarily the case. In this research, the test-takers were deliberately asked to evaluate the speaker in terms of their style of speech. This decision was made on the basis of Groelanders and Hout's (2010) research results and the fact that indirect assessment cannot be expected to work in the same way when the test-takers see the speaker as when they only hear them. If the test-takers are not told to pay attention to the style of speech and only asked to evaluate the speaker in general, it is quite possible that they include attitudes pertaining to the physical appearance of the speaker in their evaluation.

The questionnaire also included three questions on the test-takers' point of view on the comprehensibility, speech rate of the speaker and usefulness of the video. These questions also made use of the seven-graded bi-polar scale. The subjective evaluations made by test-takers could be compared to the objective intelligibility scores of the test, calculated speech rates of the speakers and differences between the control and experimental groups in test performance. These questions were used to elicit information on the informants' attitudes that may affect their performance in listening comprehension.

5.4 The sample and conducting the test

The test was conducted to two Finnish upper secondary school EFL groups in the Helsingin Normaalilyseo school in Helsinki, Finland, in January 2016; one group of first-graders and one group of third-graders. The group of first-graders consisted of 27

students 19 of whom were female and 8 were male. The mean age of this group was 16.04 years. The group of third-graders consisted of 12 students 11 of whom were female and one was male. All of the third-graders were 18 years old. Thus, the sample consisted of 39 students altogether. Conventionally, students have participated in formal English teaching for seven years when they are on the first grade of upper secondary school and for nine years when they are on the third grade. Because of the small sample size and rather uneven distribution of males and females, this research should be considered a case study. The nature of the research results is quantitative since they were elicited with a multiple choice (MC) task and a bi-polar scale speaker evaluation (SE) questionnaire.

The instructions for the test were given in English because both of the teachers whose student groups were used for the purposes of this research give all their classes in full English. However, the same instructions were written to the test papers in Finnish to make sure that everybody understood them completely. In addition, the test-takers were asked if there was anything in need of clarification before starting to play the texts. The test-takers' prior knowledge was stimulated by telling them that all of the used texts were stories. This enabled the informants to prepare for the texts' sometimes unbelievable events by activating their prior knowledge about the world of stories. Pretext stimulation is considered essential in understanding foreign languages (Dunkel 1986, Baltova 1994). This is because language seldom is without context, especially authentic language use. Therefore, the test-takers should not be thrown into the middle of a communication situation without letting them know what it is all about, otherwise it will take them a while to figure out what kind of situation it is, thus reducing their attention from the text itself.

The texts were played only once so that the test-taking situation would resemble an authentic live storytelling situation as closely as possible. Ten questions were asked of each text. 30 second breaks were taken after each stretch of about one minute of the text and two questions were answered during each break. After each text, the test-takers had two minutes time to answer the last two questions and the nine questions in the attitude questionnaire as well as prepare themselves for the first questions of the next text. The order in which the accents were played was as follows: Received Pronunciation, Zimbabwean accent, Japanese accent and Irish County Kerry accent. The test was started with the accent that was expected to be the easiest for the

test-takers to cope with so that they would be warmed up for the potentially more difficult accents.

Since the two groups that were used in this research were of different age and had a differing level of education in English, their results in the MC test were not intercomparable. Thus, each group had to be divided in half to get an audio-only control group and an experimental video group. Control and experimental groups were needed so that comparisons between the two conditions could be made when analyzing the test results. These halves of each group took turns in being the control group and the experimental group so that both halves would see two videos and only hear the audio of two videos. Half of the group were asked to turn around so that they would not face the screen on which the video was displayed and after each video both halves turned around so that the group which first faced the video would not face the next video.

This rather peculiar measure was taken for three reasons. First of all, the groups were so small that chance played a major role in that the language proficiencies of the control and experimental groups could have been very different if the groups were only divided in permanent control and experimental groups. If one group would get significantly better results from all of the texts, it could be due to differences in the language proficiencies of the groups just as well as differences in the listening conditions, i.e. audio-only condition and audiovisual condition. Now that both of the group halves functioned both as experimental and control groups, this difference in language proficiencies could be detected if one half would always get better results even if not being able to see the video.

The second reason was motivation which plays a major role in listening comprehension. The control group could become uncooperative if they knew that their classmates were able to see all of the videos when they could not. As Baltova's (1994) research results show, students find the same texts significantly more appealing when they see the video and not only hear the audio – a conclusion with which also Dahl and Ludvigsen (2014, 828) agree in a much more recent study.

The third reason was attention. According to Baltova's (1994) research results, the visual mode helps students concentrate for a longer period of time since they do not have to focus on one mode only, i.e. the sound, all the time, which burdens the capacity of their working memory more than dividing their attention between two modes.

At this point, a notion about the validity of this research can be made. As explained above, the procedure of dividing the age groups into two halves that took turns in being the video and audio-only group was conducted in order to find out if there are significant differences in the English listening comprehension proficiencies of these group halves. As it turned out, neither group half dominated the results in all texts in neither age group. Thus, it can be stated that the differences in the proficiencies of the group halves were not great enough to distort the results of this research.

Since the two student groups used in this research were not comparable to each other, the MC test was made harder for the group of third-graders that were assumed to be more proficient in English because of their longer education in the language. This was done by adding correspondence between the answer options of five questions that were found out to be the easiest ones for the group of first-graders in each text, but carefully avoiding ambiguity. In addition, correspondence between the wrong answer options and the text was added. As Yanagawa & Green point out, this is a strategy with which the difficulty of a listening comprehension test can be increased (Yanagawa & Green 2008, 112). Another reason for making the test harder was that the first-graders performed better than expected in the text which could be due to the fact that the Helsingin Normaalilyseo school frequently makes it to the top of in comparisons between the Matriculation examination results of Finnish upper secondary schools. So, the students who participated in this research were quite likely more proficient in English than the average student in their age group. The difficulty of the test was increased to ensure that there would be variety in the test results of the older group, as well.

Of course, the conditions in which the test was conducted were not ideal for the purposes of this research. If the informants of this research had been two groups of same age and roughly same size, this research could have been conducted much more traditionally. These ideal conditions were actually the conditions for which the test was originally developed. However, alterations had to be made because the groups whose time and participation were granted for this study were not of same age nor size. If the ideal conditions had been met, one of the groups would have been used as an audio-only control group and one group as an audiovisual experimental group. Thus, the structure of the research would have been simpler and easier to interpret. This having been said, the methodology can be put behind and focus shifted towards the results.

6 Results and discussion

This section is dedicated to both results and discussion. As there were three major variables, namely visual mode, accent and attitude, it is more fitting to discuss the results right after presenting them rather than first presenting all results and coming back to them later in a separate discussion section. This way it is easier to pay attention to both the results presented as figures or tables and the following discussion which aims at explaining why these results were gained. This section will quite closely follow the structure already familiar from the theory section: first the effects of the visual mode will be discussed, secondly the effects of accent, then individual questions which produced results that stood out will be discussed and finally attention will be paid to the results of the attitude questionnaire. The interrelationship between the three major variables will be discussed throughout this section when relevant. The acronyms RP, ZIM, JPN and IRE will be used to refer to the accents in the videos; Received Pronunciation, Zimbabwean accent, Japanese accent and Irish Co. Kerry accent respectively.

6.1 The effect of the visual mode

This section will explore the effects of the visual mode that were found when observing the results of the MC test. The general effects will be dealt with first by paying attention to the mean scores of each group in the MC test. After that, more detailed findings will be discussed. This will also help explain the general tendencies in the results of the MC test.

6.1.1 General effects of the visual mode

As speculated in the theory section, being able to see the video was found to facilitate intelligibility most of the time. The video group outperformed the audio-only group in six cases out of eight. However, the deviating two cases in particular show that the visual mode may distract EFL learners, as well. In both age groups, the video group outperformed the audio-only group in three out of four texts. These texts were texts 1 (RP), 2 (ZIM) and 4 (IRE) for the 1st grade group and texts 1 (RP), 3 (JPN) and 4 (IRE)

for the 3rd grade group. The greatest differences between the video and control groups appeared in the 1st grade group's results of Text 2 (ZIM) and the 3rd grade group's results of texts 2 (ZIM), 3 (JPN) and 4 (IRE). In all of the 3rd graders' results mentioned, the difference in the mean scores of video and audio-only groups was more than one point (out of the theoretical maximum score of 10 points). These findings are illustrated in Table 1 which shows the mean scores of each group in each text and the difference between the video and audio-only groups. The difference has been calculated by subtracting the mean score of the audio-only group from that of the video group.

The difference of the two conditions	Text 1: RP	Text 2: ZIM	Text 3: JPN	Text 4: IRE
1 st grade video group	9.57	8.31	9.00	8.62
1 st grade audio-only group	9.38	7.57	9.23	8.14
1st grade group difference	0.19	0.74	-0.23	0.48
3 rd grade video group	8.50	7.33	9.00	8.50
3 rd grade audio-only group	8.17	8.50	7.83	7.33
3rd grade group difference	0.33	-1.17	1.17	1.17

Table 1. Results of the MC test and the differences between groups. Negative values mark instances in which the audio-only group outperformed the video group. Theoretical maximum score in the MC test was 10.00 points.

A further observation can be made: as the task was made more difficult, the importance of the visual mode became more significant; the difference between the performances of the video and audio-only groups were greater in three out of four texts in the 3rd grade group's results when compared to the results of the 1st grade group's results. However, the group sizes also played a role; small differences in the performance led to greater differences in the mean scores when it came to the 3rd grade group which consisted of only 12 informants. When a more thorough understanding of the texts was required, the test-takers relied more on the visual mode.

Making those questions that were found to be easy for the 1st grade group harder for the 3rd graders seemed to make the questions that originally were difficult even more difficult. This can be seen in the 3rd grade group getting lower scores in many questions that both age groups had. This is probably due to two things: the text was harder to follow because less help could be gained from the questions and answers and the more difficult questions burdened the test-takers working memories more. It is also possible that the questions did not pace the text as much as for the 1st graders;

if the 3rd graders did not know for sure when they had heard the correct answer in the text, they had to focus on two questions simultaneously which obviously places great burden on the working memory.

The surprising results, namely the 1st grade audio-only group's outperformance of the video group in Text 3 (JPN), the 3rd grade audio-only group's outperformance of the video group in Text 2 (ZIM) and the 3rd grade video group's better performance in Text 3 (JPN) in comparison to Text 1 (RP) will be explained in section 6.4 which deals with the results of individual questions. The effect of accent and the interrelationship of the visual and audial modes will also be discussed while explaining what made these individual questions either so easy or difficult.

6.1.2 The effect of visual mode on the intelligibility of accents

The differences between the intelligibility of accents were only very slightly greater when the video was not seen. This suggests that generally, seeing the video does not help even the differences between the intelligibilities of accents. This can be seen when observing the mean differences between the intelligibility of accents. The mean difference was found to be only very slightly greater in the results of the audio-only groups of both age groups. The mean difference between the intelligibility of the accents of the 1st grade video group was 0.93 points whereas the same difference in the case of the audio-only group was 1.06 points. The difference of the mean scores of these groups is thus 0.13 points. The mean scores of the 3rd grade video and audio-only groups were 0.22 and 0.28 respectively. Thus, the difference of the mean scores is only 0.06 points. A conclusion that the effect of accent is very similar in both audiovisual and audial conditions can be made.

6.2 Detailed findings of the visual mode

In this section, findings of more detailed aspects of the visual mode will be discussed. The section mainly focuses on the general effects of lip-reading and gestures. However, all findings of these aspects cannot be fully covered without explaining the results of individual questions. Therefore, these topics will be discussed in even more detail in section 6.4.

Unlike in the results of Sueyoshi and Hardison (2005), no clear benefit could be found from lip-reading. The visual mode was not found to be useful if no pronounced gestures were present. This may be because the movement of lips must also deviate when the pronunciation deviates; if the listener does not understand a word because of unfamiliar pronunciation, they do not recognize the movement of lips either because different kind of movement of lips is required to produce the sounds that deviate from the familiar pronunciation. Summing this finding up, the movement of lips is not readable if the pronunciation is unfamiliar. The finding that EFL learners do not benefit from lip-reading when listening and watching accented speech can be further supported by comparing the two least intelligible speakers used in this test, the Irish and Zimbabwean speakers. The Irish speaker whose lips are not visible because his long mustache covers his mouth was more intelligible for both age groups than the Zimbabwean speaker whose lips are clearly visible. Even if lip-reading did not seem to be helpful when trying to understand accented speech, it may still be useful when listening to a familiar accent. Thus, this finding is not contrary to Sueyoshi and Hardison's (2005) results.

The visual mode was found to be particularly facilitative when clear and pronounced gestures were used as an aid to tell the story. This could be done by placing emphasis using beat or metaphoric gestures to a passage in which the correct answer to the question appeared or by using iconic gestures that clearly described a physical action or a concrete object that was relevant to the story. Beneficial emphasis could also be placed by other types of body-language than hand-arm gestures, for example by shaking one's head. This finding was made by observing questions in which the speaker's accent caused difficulties for the test-takers, but the video group performed clearly better than the audio-only group and clear gestures were present in the text.

The video groups clearly constructed meaning for the text by combining both audial and visual cues. However, the speaker's gestures indeed had to be pronounced and their meaning clear in order to be helpful; subtle gestures and gestures that were partially out of the picture had no effect. This finding was made when scrutinizing the results of questions in which the accent clearly caused difficulties but no difference was found in the results of video and audio-only groups.

The visual mode was also found to be harmful in some cases. The test-takers often seemed to rely on the visual cues for good or worse; they could be misled by ambiguous iconic gestures even when the accent was familiar, i.e. RP. The test-takers

relied on the speaker's gestures when trying to determine what the main ideas of the text were; they seemed to assume that the correct answers appear when the speaker's gestures are most pronounced. Therefore, also the beat and metaphoric gestures used to place emphasis on the text could be harmful. When the questions were not about the main ideas of the text but rather of more detailed aspects of it, many informants chose answer options that resembled the text in utterances in which the main ideas of the text were expressed. This finding was based on questions in which the audio-only group outperformed the video group and pronounced gestures were present in the text.

6.3 The effect of accent

The effect of accent was inspected by comparing the results of the MC test of each text within each group. The score of Received Pronunciation (RP) was chosen as the base to which the scores of all other accents were compared since it definitely was the accent that the test-takers were most familiar with. This is due to RP being the most used variety in listening comprehension and textbook texts as well as the pronunciation model in the English text book series used at the Helsingin normaalilyseo school. This can be seen from the results of my previous research that studied the accents used in the Open Road text book series (Iivanainen 2013). Therefore, RP can be considered a neutral accent; an accent which does not have a hindering effect on listening comprehension. As was explained in the theory section of this study, experience is the most important variable that determines the intelligibility of an accent for language learners (Smith & Bisazza 1982, Gass & Varonis 1984, Pihko 1994, Pihko 1997, Bradlow & Bent 2003a, Bradlow & Bent 2008, Kennedy & Trofimovich 2008, Adank et al. 2009). The results of the MC test are shown in Table 2.

MC listening comprehension test	Text 1: RP	Text 2: ZIM	Text 3: JPN	Text 4: IRE
1 st grade video	9.57	8.31	9.00	8.62
1 st grade audio-only	9.38	7.57	9.23	8.14
3 rd grade video	8.50	7.33	9.00	8.50
3 rd grade audio-only	8.17	8.50	7.83	7.33

Table 2. Results of the MC test displayed as mean scores of the groups. RP stands for Received Pronunciation, ZIM for Zimbabwean accent, JPN for Japanese accent and IRE for Irish accent. Theoretical maximum score in the MC test was 10.00 points.

What strikes the eye first in this table is that surprisingly, it seems like RP was not the easiest accent to understand for the 3rd grade groups. However, it has to be borne in mind that the 3rd grade groups consisted of only six informants each because of which small differences in the performances of individual students easily cause great differences in the mean scores. The reason for the 3rd grade video group to score a higher score for the Japanese speaker has actually to do with visual cues; there was one particularly tricky question devised for the more difficult MC test of the 3rd graders in which the gestures of the RP speaker misled many of the test-takers – including everyone in the video group – into choosing a wrong answer option. Had the video group performed similarly to the audio-only group in this question, the mean score of RP would have been higher than that of the Japanese accent. This question (3rd grade group Text 1, question 9) will be dealt with in more detail in section 6.4.2.

The 3rd grade audio-only group's better performance with the Zimbabwean accent than with the RP accent cannot be explained with as much certainty. A closer look at the group's performance reveals that there is a greater variation in the results of the RP text than in those of the Zimbabwean accent text. The group being so small, this fluctuation led to the Zimbabwean accent getting a higher intelligibility score. Even though it seems that the difference is 0.33 points, this difference is not very great because of the small group size. Furthermore, I remain doubtful that the Zimbabwean accent would have been easier to understand than RP for this group since the Zimbabwean accent was distinctively more difficult to understand than RP for all other groups. This anomaly can also be a product of chance; after all, guessing is possible in MC tests.

6.3.1 Differences between the intelligibility of accents

By subtracting the mean scores of each accent from the mean score of the neutral accent, RP, we can see the degree to which these accents were less intelligible than RP. Table 3 shows how great these differences in intelligibility were.

Difference between accents	ZIM	JPN	IRE
1 st grade video	1.26	0.57	0.95
1 st grade audio-only	1.81	0.15	1.24
3 rd grade video	1.17	-0.50	0.00
3 rd grade audio-only	-0.34	0.33	0.84

Table 3. The mean intelligibility scores subtracted from that of RP. The texts that gained a higher intelligibility score than the RP text show as negative values.

For the 1st grade video group, the Zimbabwean accent was clearly the most difficult one. The differences between the intelligibility of the accents were more substantial in the results of the 1st grade audio-only group. For them, the Zimbabwean accent was clearly the least intelligible one with the Irish accent the second least intelligible. The difference between these accents and RP is rather great: 1.81 points for the Zimbabwean accent and 1.24 points for the Irish accent. In this research, the reduction of one point in the mean score means that on average, one more question out of ten was answered wrong. Thus the difference of 1.81 can be considered meaningful as it means that on average about two more questions were answered wrong. The Japanese accent, however, was nearly as intelligible as RP for this group.

The results of the 3rd grade groups are much more difficult to interpret due to the reasons explained in the previous section; small differences in the performance of the students led to seemingly great differences in the mean scores. However, it is still reasonable to use mean scores for interpreting the 3rd graders' results so that consistency in the analysis persists. The Zimbabwean accent was the only one to cause any difficulties in understanding for the 3rd grade video group. For the 3rd grade audio-only group the Irish accent was distinctly the most difficult one and the Japanese accent the second most difficult but with only a 0.33-point difference to RP.

6.3.2 Interpretation of the mean scores

There appears to be no consistency on the difficulty of accents that would carry over from group to group, although the Zimbabwean accent was the most difficult one for three groups. The lack of consistency is most probably due to the fact that another variable, namely the visual mode, also had an effect on the results when it came to the video groups. However, mean values of all of the results of each age group can be calculated so that the accents can be arranged in an order from the easiest to the most

difficult one. The age groups have to be dealt with separately since the MC tests of the two age groups were different. Therefore, the groups are not intercomparable.

The mean intelligibility scores of the accents of the 1st grade groups in order from the most intelligible to the least intelligible one (with a theoretical maximum score of 10): 1. Received Pronunciation (9.48), 2. Japanese accent (9.12), 3. Irish accent (8.38), 4. Zimbabwean accent (7.94). The mean intelligibility scores of the accents for the 3rd grade groups arranged similarly: 1. Japanese accent (8.42), 2. Received Pronunciation (8.34), 3. Zimbabwean accent (7.92), 3. Irish accent (7.92).

Now, a pattern can be distinguished; RP and Japanese accents were the two most intelligible accents with only a small difference between their intelligibility scores for both age groups. Similarly, the Zimbabwean accent and Irish accent were the two most difficult accents with either a small difference between their intelligibility scores in the case of the 1st graders and with no difference in the case of the 3rd graders.

If the effect of video that distorted the results of the 3rd grade video group's performance is taken into account (see section 6.4.2: 3rd grade group Text 1, question 9), RP is the most intelligible accent for the 3rd grade groups, as well. Then the only difference in the order of difficulty between the age groups is that the Zimbabwean accent was the least intelligible for the 1st graders whereas the Zimbabwean accent and the Irish accent shared the position of the least intelligible accent for the 3rd graders.

A more detailed analysis on the difficulties the accent imposed on the test-takers' understanding of the text will be made in section 6.4. Attention will be paid on individual questions that caused difficulties for the test-takers. They will be discussed with the effects of the visual and audial modes and their interrelationship in mind.

6.3.3 The high intelligibility of the Japanese accent

There are three possible explanations for the Japanese accent to be so distinctively more intelligible than the Irish and Zimbabwean accents; possible familiarity with the accent, slow speech rate of the speaker and mismatched interlanguage speech intelligibility benefit (ISIB). Since the Japanese accent was understood with such ease, we could assume that the test-takers were familiar with the accent. As explained in the theory section, familiarity with accent is often considered the most important factor in understanding accents.

However, the conclusion about the familiarity of the Japanese accent could be too rushed; the Japanese speaker spoke clearly most slowly with a speech rate of 130 words per minute (w/m). To put that into perspective, the speech rates of the other speakers were 141 w/m by the Zimbabwean speaker, 153 w/m by the RP speaker and 199 w/m by the Irish speaker. Fast speech rate is most often a feature associated with native speakers and leaves little time for processing (Pihko 1997, 45-46) whereas slow speech rate is more typical to EFL speakers and naturally gives more time for processing the speech signal (Pihko 1994, 100-101). Thus, the speech of the Japanese speaker cannot have burdened the working memories of the informants as much as the other accents and less developed schemas were needed for decoding the message.

The slow speech rate also left room for clear and distinctive pronunciation of vowels and consonants which most likely helped the test-takers understand the accent even if it generally deviated from the norm. In other words, the slow speech rate enabled the test-takers to benefit from mismatched ISIB. As an EFL accent, the Japanese accent lacked the prosodic reduction phenomena that were present in the other accents that were either native accents, namely RP and Irish accents, or an ESL accent, the Zimbabwean accent. As explained in the theory section, the lack of prosodic reduction phenomena is the primary cause of mismatched ISIB (Bent & Bradlow 2003b, 1602).

Theoretically speaking, the higher intelligibility of the Japanese accent could have been explained by a more positive attitude towards the variety as Pihko (1997) and Major et al. (2005) regard the attitude one of the most important factors affecting the understanding of accented speech. The Japanese accent gained a mean score of 4.94 (on a scale from one to seven in which four is the neutral value) of all attitudinal aspects measured. This means that the general attitude towards the Japanese accent was much more positive than the attitude towards the Irish accent which gained a mean score of 4.06. However, the difference between the Japanese and Zimbabwean accents was only a narrow one since the Zimbabwean accent gained a mean score of 4.82, only 0.12 points less than the Japanese accent. Therefore, a positive attitude cannot be considered to be the cause of the better intelligibility of the Japanese attitude in comparison to both Irish and Zimbabwean accents. The role of attitude will be discussed more thoroughly in section 6.5.

6.3.4 The role of speech rate

As the role of speech rate was already touched upon, it is relevant to discuss it further. Slow speech rate was found to be a factor that facilitated intelligibility in the case of the Japanese accent (speech rate of 130 w/m) and fast speech rate must have affected the difficulty of the Irish accent as it was found to be the most difficult accent to understand and the Irish speaker also spoke on the fastest pace (199 w/m). However, the role of speech rate is not that simple; RP was the most intelligible accent even though it was spoken on the second fastest speech rate (153 w/m) and the Zimbabwean accent was found to be the least intelligible accent even though it was spoken on the second slowest pace (141 w/m). Furthermore, the Irish and Zimbabwean accents were as difficult for the 3rd grade group even though there was a difference of 51 w/m in the speech rates. Therefore, we can conclude that speech rate is a factor that has meaning in the intelligibility of accents but it does not dominate over other factors. Still, it indeed influences the other factors as was theorized in section 3.1.

6.3.5 Comparison to previous studies

The results of this research support the view shared by Pihko (1997) and Smith and Bisazza (1982) that the most understandable language variety for EFL learners is the standard variety that has been used most at school. However, the results cannot be considered to contradict with the view of Wilcox (1978) and Moinzadeh (2012) that the speaker's own variety is the most intelligible one, since Finnish accent was not included in the test.

The results of this study parallel those of Pihko's (1997) research; the most familiar variety was found to be the easiest one to understand; RP in this study and what Pihko called 'Slow Pedagogical' which was RP spoken on a slow speech rate in Pihko's study. In both studies an EFL variety was found to be the second easiest to understand; Japanese accent in this study and German accent in Pihko's. ISIB was found to have affected the intelligibility of the EFL variety in both studies, although Pihko did not explicitly state it since the theory had not been composed yet in 1997. However, Pihko mentions the clear energetic articulation and word-by-word speaking style (Pihko 1997, 169) and a resemblance to Finnish accent (Pihko 1997, 170) as the reasons for the high intelligibility of the German accent. Both of these characteristics

are found in the ISIB theory (Bent & Bradlow 2003b). Furthermore, an African variety, namely Gambian English, and a regional native variety, Midwest American, were found to be the two least intelligible accents in Pihko's study. Likewise, an African variety and a regional native variety, Zimbabwean English and Irish Co. Kerry accent were found to be the two most difficult varieties with a clear difference to the two most intelligible accents in this research. This suggests that Finnish EFL learners were not familiar with African English varieties and regional varieties in 1997 which has not changed over the 19 years that have passed since Pihko's research.

What is surprising, is the similarity of the intelligibility gap between the most intelligible and least intelligible accents in both researches; 19.08 percentage points in Pihko's research and 18.10 percentage points in this study when measured from the bigger and therefore more reliable audio-only group, the 1st grade audio-only group. In this case, an audio-only group has to be used for the comparison, because Pihko's research did not include video. However, these percentages should be interpreted with caution since Pihko used a different method, partial dictation test, to determine the intelligibility of the accents and her subjects were 2nd graders of upper secondary school. Furthermore, the gap between the most intelligible and least intelligible accents in the 3rd grade audio-only group was distinctively narrower, only 8.40 percentage points.

The fact that the differences in the intelligibility of the accents were narrower when it came to the 3rd grade group suggests that EFL learners learn to adapt to different accents to some extent during the upper secondary school. The mean difference between the intelligibility of RP and other accents was 0.89 points in the case of the 1st grade groups and 0.25 points in the case of the 3rd grade groups. Therefore, we can assume that some progress has been made in teaching EFL learners to cope with different accents even though African accents remain the ones causing most difficulties.

6.4 Discussion of individual questions

In this section, attention is paid to those questions that were found to be difficult either to the audio-only groups or video groups or in some cases, both. These instances help explain the general findings discussed in sections 6.2 and 6.3. The relationship

between the effects of accent and visual cues or the lack of them will also be discussed. The International Pronunciation Alphabet (IPA) will be used to explain the deviating pronunciation of individual words or phrases and to compare the pronunciation to RP since it was used as the neutral accent in this research. Only the key words and phrases needed to answer the questions in the MC task will be transcribed using IPA because it is impossible to know which words exactly the test-takers found difficult to understand. However, if they failed to answer the question correctly, it is quite safe to assume that the key words were not understood. This section starts by digging into questions in which the video group was found to have benefited of seeing the visual cues. Then questions in which the visual cues were found to have distracted the video group will be dealt with. Next, questions with ambiguous results are examined and finally questions that were equally difficult to both video and audio-only groups will be discussed.

I would like to remind the reader here that even though it may appear that there is strong resemblance between the answer options and the text excerpts presented in this section, this was not the case in the actual test. The questions and answer options presented in this section are translations of the original ones that were written in Finnish. For full transcriptions of the texts see Appendix 1 and for the questions and answers in Finnish and English respectively see appendixes 2 and 3.

Furthermore, it has to be noted here that since half of the questions were different in the tests conducted for the two age groups, even the questions that both groups had cannot be considered completely comparable to each other even though paralleled frequently in this section. This is due to the fact that the questions that were different may have had an effect on the way the test-takers generally constructed meaning for the text. Unfortunately, it is hard to estimate how this effect manifested.

6.4.1 Facilitative visual mode

The 3rd grade video group clearly outperformed the audio-only group in question 3 of Text 1 (RP): “Why would Tiresias know the answer to the question?” The correct answer option c) “Because he had lived both as a man and a woman.” was chosen by 100% of the video group whereas a mere 33% of the audio-only group chose the same answer. All of the test-takers who chose another answer option chose the answer option a) “Because he had studied the mating of snakes.” The passage of the text in which the

correct answer appeared goes as follows: “And Zeus said there’s only one person who will know the answer for that question and that is Tiresias. He alone has lived both as male and as female.” The difference of 66% between the groups is so great that it is quite clear that the video group benefited from visual cues. The cues in question are the beat gestures of the speaker that place emphasis on the passage where the correct answer lies. Interestingly though, no significant difference was observed between the video and audio-only groups of 1st graders even though they had the same question (the percentages of the 1st grade groups were 86% and 92% respectively). This could suggest that the 3rd grade video group’s better performance was due to chance, but since the difference between the groups is as great as 66% this does not seem likely. Still, we have to keep in mind that the results of the 3rd grade group are more prone to the effects of chance since the group was significantly smaller.

The results of the 1st grade video group were considerably better than the audio-only group’s in the question 8 of Text 2: “How did the hunter manage to make the mouth remain ajar?” 100% of the video group chose the correct alternative whereas 71% of the audio-only group did so. The right answer was b) “He placed a rock between the teeth.” and the passage was: “So, he took a small rock, placed it between the teeth of the hippo and the mouth remained open a little bit.” The speaker uses iconic hand gestures to describe the rock and how the hunter propped it between the teeth of the hippo. This clearly helped the video group answer correctly. The most popular wrong answer was d) “He had to use all of his strength to keep the mouth ajar.” This was likely due to the speaker saying “So, he summoned all the energies and powers that he had, pushed hard and slowly the mouth opened a little bit.” However, this utterance does not correspond to the answer option in that it only explains how the mouth was opened but not how it was *kept* open. The speaker’s accent made it difficult to understand the utterance in which the answer to the question appeared; his pronunciation clearly deviates from mainstream pronunciation when he utters it. Unlike in the example above, there is no question about the trustworthiness of this finding; the 3rd grade group had the same question and their percentages for the correct answer were 100% for the video group and 83% for the audio-only group. Furthermore, the answer option d) was the most popular wrong answer among this age group, as well.

Question 7 of Text 4: “What did the priest do on the fourth day?” caused great difficulties for the 3rd grade audio-only group; only 17% of them were able to answer

correctly whereas 83% of the video group chose the correct alternative. The correct answer was c) “He held a sermon in which he called the people superstitious.” and the passage in which the answer appeared was: “So, God almighty, at mass that Sunday, he preached them a sermon of the... out of calling them all kinds of superstitious peasants and blaa blaa blaa blaa. All that kinds of stuff.” The storyteller mimics the overly dramatic hand gestures of the priest who is giving a sermon. These gestures could be considered a combination of beat, iconic and metaphoric gestures. The gestures clearly helped the video group understand that a preaching situation is in question which helped them choose the correct alternative. None of the wrong answer options had anything to do with the text. Therefore, the assumption that those who answered wrong simply tried to guess can be made. Their guesses were most likely based on background knowledge since two of the other answer options were designed so that they could be assumed to be something that a priest would do. Thus, they had to rely on top-down processing. These answer options were b) “He rested for the fourth day was Sunday.” and d) “He prayed help for the situation for the whole day.” None of the wrong answer options gained significantly more popularity than others. This suggests that the test-takers who answered wrong had no idea what the speaker was saying because of his accent. The passage was made even more difficult because of the unexpected pause the speaker held in it. As was discussed in section 3.1 of this study, illogical or unexpected pausing makes speech less intelligible (Pihko 1997, 46). No significant difference was observed in the results of the 1st grade groups; 62% of the video group and 57% of the audio-only group answered correctly. So, the difference between the groups was only five percentage points.

In question 10 of Text 4 (IRE): “What does the storyteller say about people who say that black monster dogs do not exist?” 92% of the 1st grade video group answered correctly but the percentage for the audio-only group was 64%. The correct answer was b) “They have never seen such a dog and do not know what they are talking about.” and the passage in which the answer appears goes “Anybody who will tell you that these things don’t exist doesn’t know what he’s talking about. Or that it means is he’s never met them.” What is peculiar about the speaker of the fourth text is that the movement of his lips is impossible to see because he has a long moustache that covers his mouth. Therefore, the video group could not have benefited from seeing the movement of the speaker’s lips. The speaker uses metaphoric hand gestures and shakes his head while saying the utterances, in which the correct answer appears, making them

seem more dramatic. Thus, he places emphasis on these utterances. This directed the video group's attention to the right utterances in the passage and helped them choose the correct answer option. There was no single wrong answer choice that would have been more popular than others. What is interesting, though, is that two informants did not answer the question at all. These two facts would suggest that because of difficulties caused by the accent, many of the test-takers were completely lost at this point of the text and therefore could not answer the question correctly or did not even bother to try. The 3rd grade group had the same question but 83% of both video and audio-only groups got the correct answer. Apparently the accent did not cause such great difficulties for the 3rd grade groups in this question. Therefore, no difference between the video and audio-only groups could be observed.

6.4.2 Hindering visual mode

A dramatic 0% of the 3rd grade video group got the correct answer for question 9 of Text 1: "What did Hera do to Tiresias?" The question was rather difficult for the audio-only group, as well, but a total of 67% of them answered correctly. This was one of the questions intentionally made harder for the 3rd grade age group since the original question caused no serious difficulty for the 1st graders. The question was made more difficult by adding similarity between the answer options which were: a) "She hit him on the eyes blinding him.", b) "She punctured his eyes with her fingers.", c) "She gouged his eyes out with bare hands.", and d) "She forced Tiresias to blind himself." The passage in the text was: "And she reached down to the foot of the bed and she grabbed Tiresias by the hair and with her fingers she gouged out his eyes." Thus, the correct answer is the answer option c). However, 83% of the video group chose the incorrect alternative b). The reason for this becomes obvious when observing the speaker's gestures; he makes an iconic hand gesture which describes how Hera thrust two fingers forward, but there is no gesture depicting the act of gouging out the eyes. Thus, the video group was fooled by this ambiguous gesture and the fact that the speaker mentions that Hera used her fingers. Therefore, the visual cues misled the test-takers. What is significant is that the students seemed to construct meaning more on the basis of the visual mode than the audial mode. It is also possible that they were not familiar with the expression 'to gouge out', but this seems unlikely since most of the audio-only group seemed to be familiar with the expression.

The results of this question explain why the video group seemingly understood the Japanese accent better than RP. There was no this kind of question in Text 3 (JPN) that would have dramatically lowered the mean score of the group. In other words, this one question turned the tables for the benefit of the Japanese accent. Hypothetically speaking, if the video group would have performed in this question as well as the audio-only group, the mean score of the RP accent would have been higher than that of the Japanese accent. Thus, the video group cannot be considered to have understood the Japanese accent better than RP.

The audio-only groups of both 1st grade and 3rd grade groups performed clearly better in question 2 of Text 2 (ZIM): “What did the great hunter think of the medicine man’s advice?” A mere 8% of the 1st grade video group and 23% the 3rd grade video group chose the correct alternative whereas 43% of the 1st grade audio-only group and 67% of the 3rd grade audio-only group answered correctly. Judging by the low percentages of correct answers, this question can be considered to have been a very difficult one. The passage the question was devised of was: “But the medicine man, you know, his words were not well received by the hunter because the hunter said: “The medicine man does not understand what I want. I want more.”” The correct answer option is d) “The medicine man’s advice was not good enough for him.” In all groups, both video and audio-only groups of both 1st and 3rd grade groups, the vast majority of those who answered wrong chose the answer option a) “The advice was good, but the hunter wanted more.” The answer option a) resembles the text to a great extent but it is clearly wrong because the first part of the answer option is quite the contrary to what the storyteller says; the great hunter clearly did not consider the medicine man’s advice to be good. Those who answered the question wrong most likely chose the strategy in which they match the likeness of individual words in the text and the answer options. As was discussed in section 5.2 of this research, this is a common misleading strategy chosen by test-takers in MC tests when they do not fully understand the text (Yanagawa & Green 2008, 110). What made this passage so difficult to understand was most probably the speaker’s pronunciation of the word ‘words’ and the phrase ‘well received’ (RP pronunciation¹: /wɜːdz/ and /wel ɪsiːvd/,

¹ RP pronunciation of words and phrases was retrieved from <https://www.wiktionary.org>. The characterizations of the speakers’ pronunciation are the researcher’s own IPA transcriptions.

speaker's pronunciation: /wɛ:dz/ and /wɔ:l .iisivd/) which deviated strongly from mainstream pronunciation. Furthermore, the speaker uses strong metaphoric and beat hand gestures when he explains that the hunter wanted more, whereas his use of gestures is far less pronounced when he explains that the medicine man's words were not well received. This makes it seem like the hunter's want for more is the focus of the passage which quite likely led to the video groups choosing the wrong answer option a) more often than the audio-only groups. Thus, the speaker's gestures directed the video groups' attention to focus on an utterance that did not contain the correct answer.

83% of the 3rd grade audio-only group answered the question 5 of Text 2 (ZIM) correct whereas 50% of the video group got it right. The question was: "How were the hippos like?" There were two popular wrong answers; a) "Docile." and c) "People-friendly." but the correct answer was b) "Scary." The passage in which the correct answer appears goes: "They really looked intimidating and formidable [...]." Those test-takers who chose a wrong alternative probably either did not make out the words "intimidating" and "formidable" because of the speaker's deviating pronunciation of them (RP pronunciation of the words: /ɪntɪmɪdeɪtɪŋ/ and /fɔːmɪdəbəl/, speaker's pronunciation: /ɪntɪmɪdeɪtɪŋ/ and /fɔːmɪdabɔː/) or they simply were not familiar with these words. The role of world knowledge in this instance is ambiguous as hippos are considered the most dangerous animals in Africa but in stories they often appear as friendly creatures. As the storyteller speaks the abovementioned passage, he raises his eyebrows distinctively. This could have been interpreted as a sign of surprise, which could make the test-takers think that there was something surprising about the hippos. So, if the test-takers paid attention to the speaker's body-language and based their answer on the interpretation that he acted surprised, they could have chosen either the right or a wrong answer; if they considered hippos to be friendly creatures of stories, they could have chosen answer option c) "Scary." but if they considered hippos dangerous animals, they could have chosen answer option a) "Docile." or c) "People-friendly." If this was the case, we can assume that the test-takers were misled by top-down processing. As the potential effect of the visual cues was twofold, the visual mode cannot be considered to have been neither beneficial nor harmful in this instance. Considering that the visual mode was not found to be harmful and that there was no clear difference between the video and audio-only groups in the 1st grade group in this question (the groups scoring 64% and 69% respectively), the possibility that the

difference between the video and audio-only groups of the 3rd graders is a product of chance has to be taken into account, or even considered probable.

The two questions, questions 2 and 5 of Text 2 (ZIM), discussed above and question 10 discussed in the previous subsection, explain why the 3rd grade audio-only group performed better than the video group in Text 2. The audio-only group's better performance was a great surprise considering that Text 2 was the one in which the 1st graders benefited of visual cues the most. The 3rd grade video group was unable to make up for the benefit the audio-only group gained in question 2 unlike the 1st grade video group. This was partially due to the fact that the 1st grade video group did not perform more poorly in question 5 like the 3rd grade group. Furthermore, the audio-only group performed better than the video group in question 10. The results of the 3rd grade group in Text 2 can be a product of chance; the students who had most difficulties with the Zimbabwean accent simply happened to be in the video group. It is hard to come up with any other explanations for this surprising deviation especially since the results of the 1st grade group should be considered more trustworthy because of the larger group size.

6.4.3 Questions with ambiguous results

The 1st grade video group's performance in question 10 of Text 2 (ZIM): "How did the hunter get into the mouth of the hippo?" was significantly better than that of the audio-only group. 92% of the video group were able to choose the correct alternative d) "He crawled inside." whereas only 43% of the audio-only group chose the correct option. The passage in which the correct answer appeared was: "He took the knife from his pocket and he crawled into the mouth of the hippo." What made the text less intelligible for the audio-only group most likely was the speaker's strongly deviating pronunciation of the word 'crawled' (RP pronunciation: /kɹɔ:ld/, speaker's pronunciation: /krɔ:l:d/). On the other hand, what made the text more intelligible for the video group was the speaker's clear iconic gesture depicting the act of crawling that he performed twice during the passage. The most popular wrong answer was a) "He dragged himself along the tongue." This was probably because the speaker mentions the word 'tongue' twice in the text stretch which the question concerns: "And he really wanted to get a big chunk of the tongue and he really pushed himself further. As he kicked to push himself further to the back of the mouth so that he could get a

good grip of the tongue, he kicked the spear.” The test-takers may have matched the word tongue to the Finnish equivalent in the answer options or they did not understand why the hunter wanted to get a good grip of the tongue and assumed it was because he was dragging himself along it. An interesting finding is that these results did not recur in the results of the 3rd grade group who had the same question in their MC test. Surprisingly, their results showed opposite results: 100% of the audio-only group answered correctly but only 67% of the video group had chosen the correct answer. The deviating pronunciation seemingly did not cause difficulty for the audio-only group but for the video group it clearly did. Furthermore, they failed to interpret the speaker’s iconic gesture of crawling or they simply did not pay attention to the visual cues. This suggests that the students who had more difficulties with the Zimbabwean accent happened to be in this group. This could have happened simply because of chance since the group size was only six people.

Question 5 of Text 3 (JPN): “What had happened by the time Motoko got off the train?” was easier for the 1st grade audio-only group than for the video group with 85% and 64% of the groups choosing the correct answer option respectively. However, the 3rd grade video group performed better than the audio-only group with percentages of 83% and 50% respectively. The correct answer to this question was a) “The crowd had thinned.” and the passage in which the correct answer appeared was: “By the time I finally got off the train, the crowd had thinned a little.” What made the passage difficult to understand for the test-takers seemed to be speaker’s strongly deviating pronunciation of the word ‘thinned’ (British mainstream pronunciation: /θɪnd/, speaker’s pronunciation: /ʃɪnd/). The most popular wrong answer for both age groups was c) “The heavy book bag had started to weigh on Motoko’s legs.” This was most likely due to the speaker mentioning a heavy book bag: “I sat with my knees together with a heavy book bag on my lap.” However, the speaker does not mention that the bag would have started to weigh on her legs. The students who chose answer option c) probably matched the words they could make out in the text to the Finnish equivalents in the answer option. There were no visual cues that could have made the question easier for the video groups to answer. The video groups cannot be considered to have benefited from lip-reading since both of them did not score better than the audio-only groups. Since no clear reason for the dissimilar results of the 1st and 3rd grade groups can be found, it seems quite safe to assume that the differences between the video and audio-only groups were products of chance; those students who had most trouble with

the Japanese accent happened to be in the video group in the case of 1st graders and in the audio-only group in the case of the 3rd grade group.

Question 5 of Text 3 (JPN) was the main reason for the 1st grade audio-only group's better performance in Text 3. The fact that the video group's worse performance in question 5 was most likely due to chance and the fact that the 3rd grade video group clearly outperformed the audio-only group in this text (with mean scores of 9.00 and 7.83 respectively) together suggest that the 1st grade audio-only group was not in a beneficial position because of not being able to see the video. Instead, the audio-only group most likely consisted of students who had less trouble with the Japanese accent perhaps because of more familiarity with it than the video group.

6.4.4 Equally difficult conditions

Question 4 of Text 2 (ZIM) was found to be equally difficult for the 1st grade video and audio-only groups with correct answer percentages of 62% and 64% respectively. The question was: "When did the hunter wake up?", the passage: "Now, the hunter early in the following day woke up [...]." and the correct answer was a) "Early in the following day." The speaker's deviating pronunciation of this whole utterance, especially the word 'early' (RP pronunciation: /ɜ:li/, speaker's pronunciation: /ʔɛli/), clearly was the cause of the difficulties the groups encountered. There were no gestures to help the video group understand this utterance or to place emphasis on it. There seemed to be a difference between the performances of the same question of the 3rd grade video and audio-only groups since 100% of the video group answered correctly and 67% of the audio-only group got the right answer. However, this means that only two test-takers answered the question wrong in the audio-only group. There are two possible explanations for the 3rd grade video group's better performance; lip-reading and chance. Because lip-reading was found not to have facilitated intelligibility in this study, the latter seems quite likely. The sample size of the 1st grade group was bigger and therefore their results should be considered more trustworthy than those of the 3rd grade group. The visual mode cannot with full certainty be considered to have been helpful in this case because the video group did not perform better in both age groups.

Roughly half of the 3rd grade video and audio-only groups chose the correct answer option for the question 9 of Text 2 (ZIM): "How did the hunter manage to make the hippo's mouth remain completely open?" The passage in which the answer

appeared was: “So, he used the spear and pushed the mouth wide open, pushed so hard that finally when the mouth was fully open, he propped the mouth open with his spear [...]” Interestingly, everybody chose between the two answer options that mentioned the word *spear* even though the speaker’s pronunciation of the word ‘*spear*’ was a deviating one (RP pronunciation: /spiə/, speaker’s pronunciation: /spie:/). Therefore, it is safe to assume that the test-takers were able to cope with the deviating pronunciation of the word ‘*spear*’. The wrong alternative a) “He cranked the mouth open with his spear.” was chosen once more often than the correct alternative c) “He propped his spear between the jaws.” The important differences between these two answer options are that a) does not answer the question and the speaker does not mention ‘*cranking*’. What is even more interesting is that the speaker uses an iconic gesture in which he pushes both of his arms upwards with his palms open. So, the gesture is in contradiction with what the storyteller actually says: “[...] he used the spear and pushed the mouth wide open [...]” However, the gesture still supports choosing alternative c) more than a) because the gesture indeed is about pushing and definitely not cranking, but apparently the contradiction with what the speaker says was enough to confuse the video group. It is also possible that the word ‘*prop*’ was either not understood or recognized by those test-takers who chose the wrong answer option. In either way, the confusing gesture seemed to be of no use for the video group.

Question 10 of Text 3 (JPN): “What had happened to Motoko’s blouse?” was found to be difficult for both 1st and 3rd grade groups. In both groups the difference between the video and audio-only groups was as small as one more wrong answer in the group that performed worse. The percentages for the groups were as follows: 1st grade video group 57%, 1st grade audio-only group 62%, 3rd grade video group 67% and 3rd grade audio-only group 50%. What made the question difficult was the speaker’s deviating pronunciation in the passage: “My blouse had come untucked, so I tucked it in. I was okay, nothing happened.” The speaker’s deviating pronunciation of the word ‘*tucked*’ made her less intelligible for the test-takers (RP pronunciation: /tʌkd/, the speaker’s pronunciation: /takt/). In addition, she does not hold pauses between the words in the phrase “tucked it in” although she generally keeps short pauses between words. This may have made the phrase sound like a single word to the test-takers. It is also possible that some of the test-takers were not familiar with the verb ‘*tuck*’. The speaker makes an iconic gesture in which she tucks in an imaginary blouse, but this gesture cannot be seen completely because the speaker’s hands are

partially out of the picture during the gesture. The video group did not seem to pay attention to this gesture because of its unclear visibility since they did not clearly outperform the audio-only group in either of the age groups. The correct answer was: a) “It had become untucked.” The most popular wrong answer was d) “Nothing happened to the blouse.” This is probably due to the speaker saying “[...] nothing happened.” Those test-takers who chose the answer option d) probably had not understood the passage about the blouse and mistakenly assumed that the statement “nothing happened” had to do with the blouse. Answer options b) “The blouse had been torn.” and c) “The blouse was covered in dirt.” were also chosen by some test-takers. They probably simply guessed what could have happened to the blouse since these answer options did not relate to the text in any way.

6.5 Results of the speaker evaluation questionnaire

This section will focus on the test-takers’ attitudes towards the accents used in the texts as well as the estimated speech rates, comprehensibility of the speakers and usefulness of the visual mode. The analysis consists mainly of quantitative observations of the mean scores of the speaker evaluation (SE) questionnaire, but some individual cases will also be explored when relevant. First, the general observations of the SE questionnaire are discussed then moving on to more detailed findings of the effects of attitude on the intelligibility of accents.

6.5.1 General observations on the status and solidarity dimensions

In all of the aspects of both status and solidarity dimension, 1st grade and 3rd grade groups evaluated the accents quite similarly; all accents were ranked in the same order from the most favorably evaluated to the least favorably evaluated accent. This order was: 1. Received Pronunciation, 2. Japanese accent, 3. Zimbabwean accent and 4. Irish Co. Kerry accent. Since the groups evaluated the accents in such similar fashion, the differences between the evaluation of accents will be observed with the use of a mean score calculated from the values given to the speaker by every participant. The mean values of both age groups were also very similar. Neither group was significantly more

positive nor negative towards any accent. The mean scores of each status and solidarity related aspect of each speaker are displayed in Figure 2.

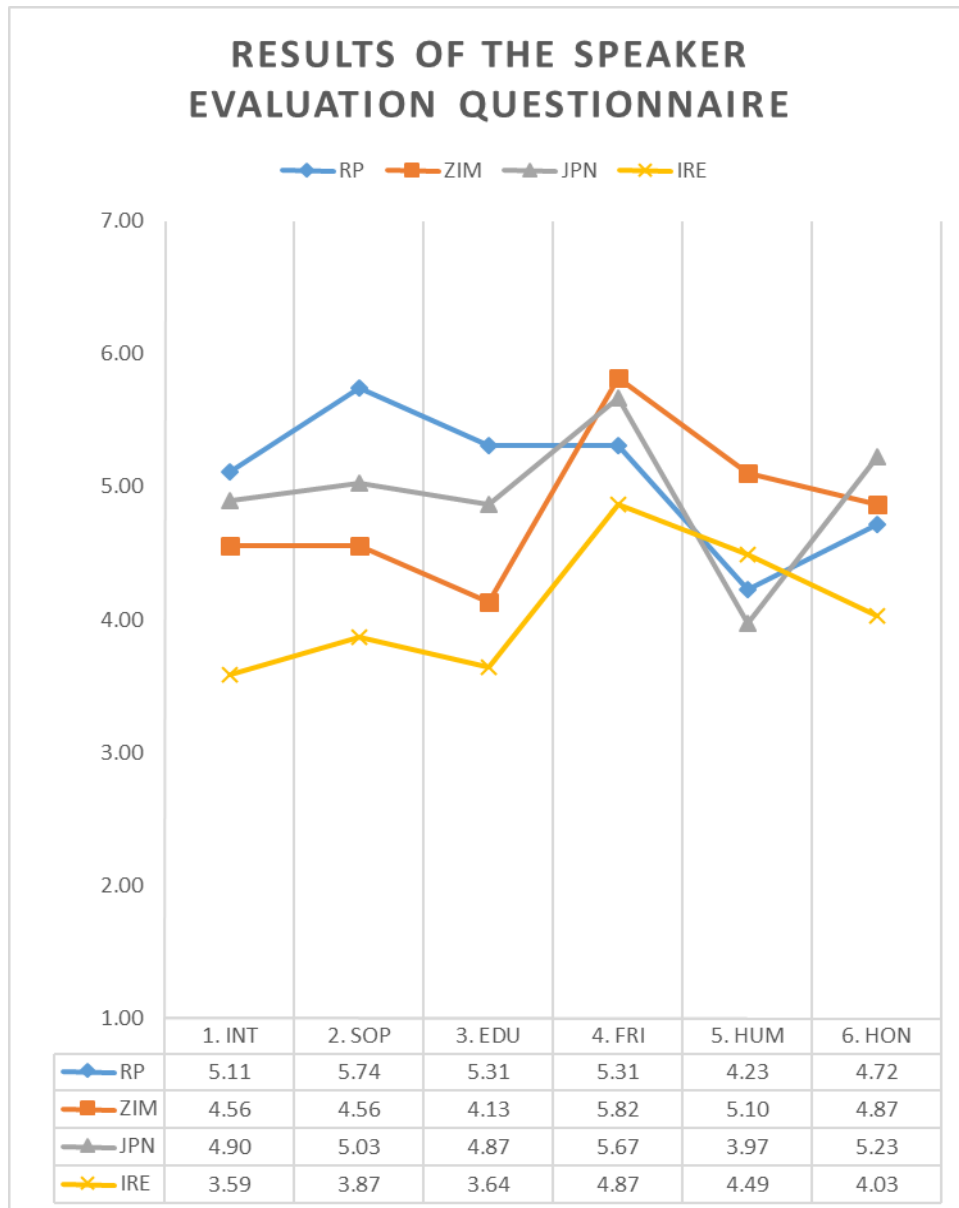


Figure 2. Mean scores in the status and solidarity dimensions. Values were given by 39 informants. RP stands for Received Pronunciation, ZIM for Zimbabwean accent, JPN for Japanese accent, IRE for Irish Co. Kerry accent. Explanations for the acronyms of the aspects: 1. INT: intelligence, 2. SOP: sophistication, 3. EDU: education. The three last acronyms stand for 4. FRI: friendliness, 5. HUM: humorousness, 6. HON: honesty. Aspects 1-3 form the status dimension and aspects 4-6 form the solidarity dimension. The values were given on a scale from one to seven in which four is the neutral value.

The results of the Speaker Evaluation (SE) questionnaire of this research follow a pattern that can be seen in previous studies: the accent used as a pronunciation model at school gained the highest score in all aspects of the status dimension; in the case of this study, this accent is RP. Similar results have been gained by various researchers (Matsuura 1994, Ladegaard 1998, Chien 2014, Rindal 2014, Buckingham 2015). Another pattern was observed by Matsuura et al. (1994), Chiba et al. (1995), Chan (2013); the difference between non-native and native accents tends to become narrower when the solidarity dimension is in question. However, in this research, the non-native varieties not only narrowed the gap between them and the school model variety but they were actually evaluated more favorably most of the time. In friendliness and honesty aspects the ESL variety, Zimbabwean accent, and the EFL variety, Japanese accent, surpassed the score of RP. In the humorousness aspect the Zimbabwean accent gained a higher score than RP whereas the Japanese accent was given a lower score.

The Irish accent is clearly conspicuous in the status dimension; it is the only accent to have gained negative scores and it did so in every aspect of this dimension. Furthermore, the Irish speaker was given the lowest score in two out of three solidarity aspects that were friendliness and honesty but in the humorousness aspect it was rated the second highest. In the solidarity dimension the Irish speaker no more gained a negative mean score, but a positive one in friendliness and humorousness aspects and a very neutral one in the honesty aspect. The mean score of all the aspects in the solidarity dimension puts the Irish speaker on the fourth position in the solidarity dimension, as well.

The Zimbabwean speaker gained slightly positive mean scores of 4.56 in intelligence and sophistication aspects and a value close to neutral, 4.13, in the education aspect. In a comparison between speakers, he was in the third position in the status dimension. In the solidarity dimension, however, the Zimbabwean speaker was clearly in the first position by gaining the highest mean scores in friendliness and humorousness aspects and the second highest score in the honesty aspect. The scores of these aspects were also very positive with a friendliness rating of nearly six, humorousness rating of over five and honesty rating of nearly five. The mean score of the aspects of the solidarity dimension puts the Zimbabwean speaker on the first position in the solidarity dimension.

The Japanese speaker was given clearly positive values in the status dimension; intelligence and education ratings of nearly five and a sophistication rating of slightly over five. In the solidarity dimension she was given very positive values in the friendliness (5.67) and honesty (5.23) aspects with the latter being the highest rating of all accents in the honesty aspect. Interestingly, she was not considered humorous despite the otherwise positive evaluation of solidarity dimension. She gained a neutral rating of 3.97 which was the lowest rating in this aspect. Judging by the mean scores of the solidarity aspects, the Japanese speaker was on the second rank in the solidarity dimension.

Received Pronunciation dominated the status dimension gaining highest mean scores in all aspects with ratings of over five in each aspect. Surprisingly, RP is on the third rank in each aspect of the solidarity dimension with the mean scores of humorousness and honesty dropping below five but still remaining on the positive side. As was discussed in section 4.3, the school model variety has been rated highest in both status and solidarity dimensions in previous researches (Matsuura et al. 1994, Ladegaard 1998, Chien 2014, Rindal 2014, Buckingham 2015). The mean scores of the solidarity dimension show that the RP speaker was on the third rank in the solidarity dimension.

6.5.2 Intelligibility and comprehensibility

In this research the terms ‘intelligibility’ and ‘comprehensibility’ are used in the way that is conventional in this type of studies; intelligibility refers to the objective intelligibility which is determined by test results and comprehensibility refers to the informants’ subjective perceived idea of intelligibility. In other words, intelligibility is the actual degree of understanding whereas comprehensibility means how well the informants think they understand the speaker.

The test-takers’ evaluation of comprehensibility followed their calculated intelligibility quite closely if only the order from the most comprehensible speaker to the least comprehensible speaker is compared to the order of the most intelligible speaker to the least intelligible speaker. On the bi-polar scale from 1 to 7, in which 1 means easy to understand, 7 means difficult to understand and 4 is a neutral value, the comprehensibility scores given to the speakers by the first graders were: 1) RP: 2.22, 2) JPN: 3.22, 3) ZIM: 3.96 and 4) IRE: 5.76. The mean intelligibility scores of the

accents calculated from the 1st graders' results were (with the theoretical maximum score of 10 points): 1) RP: 9.48, 2. JPN: 9.12, 3. IRE: 8.38 and 4. ZIM: 7.94. As we can see, the 1st graders evaluated the accents in the order that resembles the intelligibility order; the only difference is that they considered the Irish accent to be more difficult than the Zimbabwean accent.

When scrutinizing these mean values more closely we can see that the difference in the comprehensibility score between RP and JPN is an entire point although the difference between the intelligibility scores of these accents is merely 0.36 points (here we have to take into account that a difference of one point is much greater on a scale from one to seven than on a scale from zero to ten). Thus, the 1st graders considered the difference between their understanding of these accents greater than it actually was. The Zimbabwean accent received a score very close to neutral; the informants considered it neither hard or nor easy even though in reality, it was the most difficult accent to understand. On the other hand, they considered the Irish accent very difficult to understand with a score of 5.76 although the intelligibility scores show that it was actually much more intelligible to them than the Zimbabwean accent. The Irish accent was also the only accent to have gained a mean value of over four which is the neutral value. These findings support Derwing and Munro's (1997) finding that listeners often rate accented speech as more difficult to understand than it actually is.

When comparing the 3rd graders' evaluation of comprehensibility to actual intelligibility we can see a similar pattern. The comprehensibility scores given by the 3rd graders were: 1) RP 3.17, 2) JPN 3.33, 3) ZIM 4.08 and 4. IRE 6.00 and the calculated intelligibility scores were: 1) JPN 8.42, 2) RP 8.34, 3) ZIM 7.92, 3) IRE 7.92. Here we should bear in mind that the confusing gesture in Text 1 (RP) distorted the intelligibility scores so that the Japanese accent may seem more intelligible than RP even though it actually was not (as was discussed in more detail in section 6.4.2). The 3rd graders did not erroneously consider the difference between their understanding of RP and Japanese accents greater than it was like the 1st graders did. However, they considered the Irish accent much more difficult to understand than the Zimbabwean accent even though these accents received the same intelligibility score. Once again, the Zimbabwean accent gained a very neutral comprehensibility score of 4.08 whereas the score of the Irish accent was as high as 6.00.

What is noteworthy, is that the 3rd graders considered all of the accents slightly more difficult to understand than the 1st graders. This is rather unsurprising considering

that the test was made more difficult for the 3rd grade group. Therefore, they probably felt that they had more difficult time understanding the texts than the 1st graders. Even though the finding that a more difficult test leads the test-takers to consider the accents more difficult to understand is not surprising, the implication of this finding is meaningful; the more difficult the task is, the greater are the risks of listener irritation and formation of negative attitudes towards the accent, which according to Pihko's (1997) theory, effect intelligibility negatively.

6.5.3 The role of attitude

The evaluation of the status dimension by both age groups correlated directly with the comprehensibility score given by them; the order from the accent with the highest status score to the one with the lowest (1. RP, 2. JPN, 3. ZIM, 4. IRE) is the exact same order as the order from the most comprehensible accent to the least comprehensible one, which was the same for both age groups. Since the attitudes seemed to resemble comprehensibility more than intelligibility, it may be relevant to question the extent of significance that attitudes are often considered to have on intelligibility. Do attitudinal studies always give trustworthy information if both intelligibility and comprehensibility are not taken into account? Unfortunately, many studies do not consider the relationship between comprehensibility and intelligibility but discuss comprehensibility as a self-evident indicator of intelligibility. It may be possible that a positive attitude in the status dimension follows a positive idea of comprehensibility rather than that a positive attitude would facilitate intelligibility.

On average, the video groups evaluated the speaker slightly more positively than the audio-only groups in both status and solidarity dimensions in all but one case. This happened with both 1st grade and 3rd grade groups. This finding could suggest that a positive attitude fosters intelligibility, but since the video group was found to have performed better than the audio-only group only when clear gestures were present, the conclusion that the better performance was due to seeing the gestures, has to be made. This conclusion is further backed by the case in which the video group gave more negative scores to the speaker in each aspect of both status and solidarity dimensions than the audio-only group. This case was the 1st grade group's Text 4 (IRE). Even though having a slightly negative attitude towards the speaker (SE questionnaire mean score of 3.78), the video group performed better than the audio-

only group who had a positive attitude (SE questionnaire mean score of 4.42). This finding suggests that even if a negative attitude would impose a negative effect on intelligibility, this effect could not override the positive effect that seeing the video has. In addition, in the two exceptions in which the audio-only group outperformed the video group, 1st grade group Text 3 (JPN) and 3rd grade group Text 2 (ZIM), the overall attitude of video group was clearly more positive than that of the audio-only group. This suggests that the positive attitude did not facilitate intelligibility enough for the video groups to be able to perform better in the MC test.

As a conclusion to the role that attitude played in this study we can thus safely say either that the attitude did not have any effect on intelligibility or that the effect that attitude had was so subtle that it could not be observed in the results of the MC test. However, this does not mean that attitude does not matter; it may play a much bigger role in other types of communication situations, especially interactive ones. When the listener has a chance to participate in the communication situation for example by interrupting the speaker, the participants of the communication situation can construct meaning together. Then it would seem essential that the participants of the communication situation are eager to understand each other.

6.5.4 Listener irritation

One very clear case of what Pihko (1994) calls listener irritation can be found from the research data: an informant who had given generally positive scores to three of the speakers in the speaker evaluation (SE) questionnaire gave the Irish speaker the value 1, the most negative value, in every aspect that was asked. This informant considered all other speakers easily intelligible giving them values of either 1 or 2 in the ‘Difficulty of Understanding’ (DOU) item but for the Irish speaker he gave 7, the most extreme negative value. What is peculiar is that the actual intelligibility of the Irish speaker for this informant was 9/10. This finding further supports the conclusion that attitude may not be as significant factor of intelligibility as has been previously thought. However, this finding supports Derwing and Munro’s (1997, 12) theory about accented speech requiring more effort to process which easily causes the listener to consider the speaker difficult to understand even though they actually are perfectly intelligible. Furthermore, this finding supports Lopez-Soto and Barrera-Pardo’s (2007, 1603) idea of language learners’ attitudes being more negative towards varieties they

find difficult to understand. In this case, however, listener irritation did not lead to communication breakdown or unwillingness of trying to understand the speaker which according to Pihko (1997, 125) are the dangers of listener irritation.

Interestingly, the intelligibility of the Zimbabwean speaker for the same informant was slightly poorer, 8/10, but he was given a DOU value of 2 and very positive scores in the SE questionnaire – a maximum 7 in all aspects of the solidarity dimension – by the informant. This shows that the negative attitude indeed had to do with perceived comprehensibility rather than actual intelligibility. Unfortunately, the informant had for some reason, perhaps by accident, left the ‘usefulness of video’ (UOF) item unanswered in Text 4 (IRE) even though he was in the video group of this text. In Text 2 (ZIM), he considered the video completely useless giving it a UOF value of 1 which could suggest that he did not find the video useful in Text 4 either, but without certain knowledge this remains only a matter of speculation.

Generally, other cases in which the test-takers scored high in the multiple choice (MC) test but gave the speakers negative values in the SE questionnaire were found but they were not as distinctive as the case dealt with in the paragraphs above. On the other hand, also cases in which the informant found the speaker difficult to understand but still gave them positive values in the SE questionnaire were encountered. So, although support was found for Derwing and Munro’s (1997) and Lopez-Soto and Barrera-Pardo’s (2007) theories, the phenomenon is not as clear-cut as they may suggest; the theories apply in general but deviations are not all that rare. Indeed, when observing the research data on a general level, a connection between perceived difficulty of understanding and negative attitudes can be found. The less comprehensible the test-takers considered the speaker, the more negative their attitudes were towards them. All in all, these research results suggest that extreme cases of listener irritation are not very usual since only one such case could be identified from the sample of 39 informants evaluating four speakers, i.e. altogether 156 possible cases. It must be stated, however, that more cases could have been identified if the informants were for example interviewed, which is what Pihko (1997) did. The combination of MC test and bi-polar SE questionnaire that was used in this research does not give insight to the informants’ reasoning behind their choices. Therefore, listener irritation could not be explored as widely as Pihko (1997) was able to do with her interviews.

6.5.5 Estimated speech rate

The estimated speech rates of the video and the audio-only groups were very similar and there was no consistency in which group estimated the speech rate to be faster. Therefore, the mean scores can be used when making observations of speech rate and estimated speech rate. The calculated speech rates of the speakers in order from the fastest to the slowest are: 1. IRE 199 words per minute (w/m) 2. RP 153 w/m, 3. ZIM 141 w/m and 4. JAP 130 w/m. Figure 6 shows the estimated speech rates by each group and the mean scores given by each age group.

Estimated speech rate	RP	ZIM	JPN	IRE
1 st grade video	3.93	4.00	3.86	5.31
1 st grade audio-only	4.15	4.00	3.85	5.21
1st grade mean	4.04	4.00	3.86	5.26
3 rd grade video	4.00	3.83	3.83	5.40
3 rd grade audio-only	3.67	3.83	4.00	5.17
3rd grade mean	3.84	3.83	3.92	5.29

Figure 6. Estimated speech rates by each group and mean scores of each age group. Values were given on a scale from one to seven, where one stood for too slow, four was a neutral value and seven stood for too fast.

Actual speech rates and estimated speech rates were found to match very closely when observing their order from the fastest to the slowest; the 1st grade group put the accents in an order that matches the order of the actual speech rates precisely. The 3rd grade group mistakenly evaluated the JPN speaker to have spoken faster than the RP and ZIM speakers, but otherwise seemed to be able to estimate the speech rate quite accurately. When observing the estimated speech rates from another point of view which is their relation to each other, further proof of the test-takers' ability to estimate the speech rates correctly can be found. We can notice that the accents that were spoken with speech rates quite close to each other, namely RP, ZIM and JPN, were given very similar values whereas the one accent that was clearly spoken faster than others, IRE, was evaluated to have been spoken on a faster rate than the rest of the accents.

The Irish speaker was the only one who was considered to have spoken too fast; the test-takers found the speech rate of 199 w/m too fast for accented speech. Speech rates between 153 w/m and 130 w/m were found to be suitable by the informants when it came to both familiar and unfamiliar accents. This deduction can

be made since the accents that were spoken on these rates were given values very close to the neutral value four.

On the basis of their research results, Derwing and Munro (1997, 11) theorized that speech rate is one of the most important factors affecting the evaluation of comprehensibility. The results of this study, however, do not support this view; the mean scores of the evaluated speech rates did not correlate with mean comprehensibility ratings at all. The order from the most comprehensible accent (1. RP, 2. JPN, 3. IRE, 4. ZIM) to the least comprehensible accent did not match the order from the evaluated slowest speech rate to the fastest nor the actual speech rate from the slowest speech rate to the fastest. Furthermore, there was much greater variation in the mean comprehensibility values in comparison to the mean speech rate values even though they were rated on the same scale.

6.5.6 Estimated usefulness of video

To observe the relationship between the estimated usefulness of the visual mode and the actual usefulness of the visual mode, the texts need to be arranged in orders that show in which text the visual mode was most useful and in which text the visual mode was thought to be most useful. The actual usefulness of the visual mode can be seen by observing the differences in the intelligibility of the texts for the video and audio-only groups. Table 4 shows both the estimated and actual usefulness of video ratings.

The usefulness of video (UOF)	1 st	2 nd	3 rd	4 th
1 st grade actual UOF	ZIM (0.74)	IRE (0.48)	RP (0.19)	JPN (-0.23)
1 st grade estimated UOF	JPN (4.93)	RP (4.63)	IRE (2.75)	ZIM (2.54)
3 rd grade actual UOF	JPN (1.17)	IRE (1.17)	RP (0.33)	ZIM (-1.17)
3 rd grade estimated UOF	RP (4.83)	JPN (4.33)	ZIM (4.17)	IRE (3.83)

Table 4. The actual and estimated usefulness of video ratings (in order from the most useful to the least useful). The value in brackets in the actual usefulness ratings is the subtraction of the mean intelligibility score of the audio-only group from the mean intelligibility score of the video group. The value in brackets in the estimated usefulness ratings is the mean score on a scale from one to seven in which one stands for useless, seven stands for useful and four is the neutral value.

As we can see in Table 4, the actual usefulness of the video was exactly contrary to the estimated usefulness of the video when it comes to the 1st grade group.

The 3rd grade group did not estimate the usefulness of the video much better than the 1st grade group either. It seems like the informants put the intelligibility of each text on account of the visual mode rather than the accent, since the accents that were found more easily intelligible (RP and JPN) gained higher scores than the accents that were more difficult to understand (IRE and ZIM). This can be seen especially in the 1st graders' estimated usefulness of video ratings: JPN and RP have been given positive ratings whereas IRE and ZIM have been given negative ratings even though the visual mode was actually most useful in the texts with IRE and ZIM speakers and least useful in the texts with RP and JPN speakers. In other words, the informants did not consider the video useful when it actually was most useful and they considered the video useful when it actually was either least useful or not useful at all.

A noteworthy discovery about the informants' evaluation of the usefulness of video is that the groups who started as the video group clearly evaluated the video more as useful than the groups that started as an audio-only group. This is most probably due to the informants comparing the texts to each other; the groups who started as the video groups were able to see the speakers whose accents were found to be most intelligible (RP and JPN) whereas the groups who started as audio-only groups were able to see the speakers whose accents were found to be least intelligible (ZIM and IRE). The groups, who started as video groups, found the texts, in which they were able to see the speaker, easier to understand and therefore gave positive usefulness of video ratings. On the other hand, the groups, who started as audio-only groups, found the texts, in which they were able to see the speaker, more difficult to understand and therefore gave negative usefulness of video ratings. If the same informants would have been able to estimate the usefulness of video in all of the texts, the ratings could be different, but since this was not the case, it remains a matter of speculation.

7 Summary and conclusions

This research dealt with audiovisual EFL listening comprehension focusing on three major variables: the effect of the visual mode, the effect of the speaker's accent and the effect of the listener's attitude towards the speaker's accent. The research was carried out using a multiple choice (MC) listening comprehension task and a bi-polar scale speaker evaluation (SE) questionnaire. The participants were 39 Finnish upper secondary school EFL learners. Four audiovisual texts with different accents were used as the material of the MC task. These accents were Received Pronunciation (RP), Zimbabwean native Shona speaker accent (ZIM), Japanese accent (JPN) and Irish County Kerry accent (IRE).

The visual mode was found to facilitate intelligibility in listening comprehension; the experimental video groups systematically outperformed the audio-only control groups. However, the visual mode was found useful only when the speaker used clear gestures to help the perceivers construct meaning through both audial and visual modes. Lip-reading was not found to facilitate intelligibility since the visual mode was only found to have an effect on intelligibility when gestures were present. The reason for this was considered to be the unfamiliar movement of lips that is required to produce the words spoken in an unfamiliar accent. Even though the visual mode was found generally helpful for EFL students, they were also sometimes misled into choosing wrong answer options because of ambiguous gestures or because the speakers gestures placed emphasis on parts of the text that did not include the answer to the question. The gestures that appeared on the texts used in this research could be categorized as beat, metaphoric or iconic gestures. All of these gesture types were found to be crucial in constructing meaning. EFL learners were found to construct meaning more on the basis of the visual mode than the audial mode. This was concluded since they could be misled by ambiguous gestures even when the accent was very familiar, i.e. Received Pronunciation. Furthermore, the results of the SE questionnaire's 'usefulness of video' item showed that the test-takers considered the differences in the intelligibility of the texts to be due to the video.

Accent was found to have a clear effect on the intelligibility of the speaker. The greatest difference that was measured between the mean intelligibility of the familiar accent, RP, and another accent, ZIM, was 1.81 points out of the maximum of 10 points. This suggests that according to the results of this study, accent can cause a

decrease of nearly one fifth in the intelligibility of a speaker of an unfamiliar accent. The most intelligible accent was found to be RP, the second most intelligible JPN, the third most intelligible IRE and the least intelligible ZIM. A completely new finding about the relation between accent and visual mode was that the visual mode was not found to have evened the differences between the intelligibility of accents; the differences between the mean intelligibility of different accents were very similar in both video and audio-only conditions.

Another new finding is that the speakers are evaluated more positively in the SE questionnaire when they are seen by the informants; the values given to the speaker in each aspect by the video groups were systematically more positive than those given by the audio-only groups. A correlation between positive attitude and comprehensibility rather than intelligibility was found. This finding together with the finding that the visual mode was not useful when clear gestures were not present suggest that the attitude towards the speaker's accent is meaningful only in terms of comprehensibility and not in terms of intelligibility. Comprehensibility refers to the informants' subjective evaluation of intelligibility and intelligibility refers to the actual intelligibility that was calculated from the results of the MC task. Intelligibility and comprehensibility were not found to match.

The school model variety, RP, dominated the status dimension of the SE questionnaire which is a tendency that can be seen in previous studies, as well. However, unlike in previous studies, the school model variety was not evaluated most positively in the solidarity dimension. Whether this has to do with the visual mode, which has not been taken into account in previous attitudinal studies, or with characteristics specific to the sample remains unclear. The informants were able to estimate the speech rates of the speakers quite correctly which can be seen by observing the order in which the mean value of the speech rate item of the SE questionnaire organizes the speakers. This order is the same as the order in which actual calculated speech rates put the speakers. On the other hand, the informants estimated the usefulness of the video (UOF) completely wrong by giving those videos, in which the visual mode was found to be most useful, the lowest UOF rating and vice versa. Listener irritation was found to be a rare phenomenon among this sample; only one clear case could be distinguished. Unlike the current theory suggests, listener irritation did not have a strong negative effect on intelligibility. The informant performed very well despite the apparent negative feelings the speaker accent aroused

in them. However, since only one such case was encountered, this finding cannot be generalized.

Although this research has given new research information we have to bear in mind that this study has its limitations. First of all, a rather small sample of 39 students with an uneven gender distribution was studied. Therefore, the results of this research should be considered with the caution that is often advisable for interpreting the results of a case study; generalizations should not be too eagerly made. The same rule that applies to all intelligibility tests applies to the MC task used in this research: “We cannot test intelligibility without including the effect of the test itself,” (Voegelin and Harris 1952, cited in Pihko 1997, 206). The MC task was found very suitable for the purposes of this study for it did not confuse the test-takers and it could be completely objectively evaluated by the researcher. However, there is a chance that the small differences in the quality of the texts have had an effect on the test results; as the videos were not filmed for the purposes of this study but downloaded from the social media site YouTube, they were not filmed with the same equipment which obviously results in differences in quality. Furthermore, the angle of view was not the same in all of the videos although as similar videos as possible were chosen. Of course, the phrasing of the questions used in the MC task also have an effect; very different kinds of questions can be made from the same material. However, careful attention was paid when making the questions so that similar kinds of questions would be asked of each text. The speakers’ personal styles of speech could also have had an effect. As Wilcox puts it: “[...] it is next to impossible to find a truly representative speaker of any accent,” (Wilcox 1978, 8). Using different speakers of the same accents could produce differences in the intelligibility results of the accents.

What can be said about the validity of the results of the SE questionnaire is that it is possible that the content of the stories the speakers told have had an effect on the evaluation of the solidarity dimension. For example, if the story itself was humorous, the speaker could be considered humorous, or if the story was completely unbelievable and told as if it was true, the speaker could be considered dishonest. Indeed, the test and speaker may affect SE in a fashion similar to intelligibility tests. The speaker’s tone of voice or just the general sound of their voice could have an effect on SE; a neutral yet truly representative speaker of an accent does not exist. In this research also the physical appearance of the speaker and their body-language and expressions quite likely had an effect; the video groups tended to evaluate the speakers more positively

than the audio-only groups. The greatest limitation for the SE questionnaire was the lack of a tested method for conducting SE questionnaires in an audiovisual context. This could also be considered an advantage because this research produced new information on conducting audiovisual SE questionnaires.

As one inspiration for conducting this research was the possibility that audiovisual listening comprehension will be used in the Finnish Matriculation Examinations in near future, this study has a range of pedagogical implications. EFL learners seem to rely much on visual cues which should be taken into account when designing audiovisual listening comprehension tests. The visual mode can be used to make listening comprehension tests either easier or harder. Either way, much attention has to be paid to the speaker's gestures already when designing the test; the speakers gestures should be analyzed carefully because ambiguous gestures may lead to unexpected results. Furthermore, all types of gestures bore meaning in this test. Therefore, no gestures should be dismissed as non-meaningful; even if a gesture may seem peripheral, it may place emphasis on a certain stretch of text. It seems crucial that students be familiarized with audiovisual listening comprehension tests before using them in official examinations so that the effect of the test can be reduced. Thus, the student's actual understanding of the text can be tested instead of their adaptability to a possibly unfamiliar test type.

Another pedagogical implication that relates to designing listening comprehension tests is that the effect of an unfamiliar accent in listening comprehension was found to be a reduction of intelligibility up to 18% of the test's maximum points when compared to a familiar accent. If an unfamiliar accent is used in a listening comprehension test, the test designer has to take into account that the accent may greatly increase the difficulty of the test. As EFL speakers have already been used in the Matriculation Examination English listening comprehension tests, this finding surely has practical advantages. Students would greatly benefit from familiarization to many kinds of accents of English at school, both examinations and future real life communication situations in mind; like Ranta's (2010, 162) research results suggest, approximately 62% of Finnish upper secondary school students' foreign language use is English as a lingua franca. Furthermore, Louhiala-Salminen and Kankaanranta's (2012, 266) research results show that approximately 70% of the English use in global companies in Finland is between non-native speakers. Students

can therefore expect to encounter a wide variety of English accents both during and after schooling.

The implication that the results of this research have on the theory of attitudes towards accents is that positive attitudes may not foster intelligibility as much as previously has been thought. Likewise, negative attitudes and listener irritation were not found to have a visible effect either. However, this conclusion should be interpreted with caution since this research indeed is only a case study. The attitudes' lack of effect could well be due to sample related variables; it is for example possible that the informants' personal motivation overrode the effects that attitude would have had. The finding that attitude did not play an important role in listening comprehension may well apply to only the specific type of listening comprehension task used in this test, an audiovisual MC task that attempted to recreate an authentic storytelling situation. If students would have a chance to listen and watch an audiovisual text on a pace of their own choosing and use tools such as rewinding the video, it could be possible that a positive attitude could encourage the students to spend more time trying to understand the texts. Furthermore, positive attitude may well be very crucial in interactive communication situations in which meaning is constructed in cooperation between the participants. Therefore, EFL learners should be exposed to various accents so that they may develop a positive attitude and appreciation towards the great variety of English accents despite the fact that the attitude towards accents was not found crucial in terms of intelligibility in this research.

As for implications for further research, the simplest implication is to repeat a research similar to this one since it was the first study to take the visual mode, accent and attitude into account as variables of listening comprehension. If more similar researches were conducted, we could see if the results would follow a similar pattern. Now that a model for a case study on this topic has been created, also larger scale researches on the area can more easily be approached. It would be essential to research this topic with a much larger sample; the use of several EFL learner groups would allow the use of separate experimental video groups and audio-only control groups rather than two groups taking turns in being the video and audio-only group. If a larger scale research was to be funded, the texts used as intelligibility test materials could be filmed by the researcher. Thus the effect of the test could be minimized as all videos would be equal in terms of quality and angle of view. The use of actors as speakers would grant the test developer the means to control the speakers' vocabulary, speech

rates and amount and type of gestures, thus equalizing the texts in terms of speaker and story related aspects. This would, however, render the test material unauthentic since it would have been filmed specifically for the purposes of the research. This research focused only on four accents of English. Further research could also make use of accents that were not used in this study. It is also important to create a valid method for SE questionnaires in an audiovisual context as no method existed before this research and the visual mode was found to have an effect on the results of the bipolar SE questionnaire used in this research.

References

- Adank, P. et al. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology. Human Perception & Performance*. 35 (2). p. 520-529.
- Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *Canadian Modern Language Review*. 50 (3). p. 507-531.
- Baltova, I. (1999). Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. *Canadian Modern Language Review*. 56 (1). p. 31-48.
- Batty, A. O. (2015). A comparison of video- and audio-mediated listening tests with many-facet Rasch modeling and differential distractor functioning. *Language Testing*. 32 (1) p. 3-20.
- Bradac, J. J. and Giles, H. (1991). Social and educational consequences of language attitudes. *Moderna Språk*. 85 (1). p. 1-11.
- Bradlow, A. R. and Bent, T. (2003a). 'Listener adaptation to foreign-accented English.' Proceedings of the 15th International Congress of Phonetic Sciences, Evanston, Northwestern University January 2003. [Online] Northwestern University, Department of Linguistics. Available at: http://www.researchgate.net/publication/228754213_Listener_adaptation_to_foreignaccented_English. [Accessed: 25th August 2015].
- Bradlow, A. R. and Bent, T. (2003b). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*. 114 (3). p. 1600-1610.
- Bradlow, A. R. and Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*. 106 (2). p. 707-729.
- Bresnahan, M. J. et al. (2002). Attitudinal and affective response toward accented English. *Language & Communication*. 22 (2). p. 171-185.
- Buckingham, L. (2015). Recognising English accents in the community: Omani students' accent preferences and perceptions of nativeness. *Journal of Multilingual & Multicultural Development*. 36 (2). p. 182-197.

- Cargile, A. C. and Giles, H. (1997). Understanding Language Attitudes: Exploring Listener Affect and Identity. *Language and communication*. 17 (3). p. 195-217.
- Chan, J. Y. H. (2013). Contextual variation and Hong Kong English. *World Englishes*. 32 (1). p. 54-74.
- Cheng, H. (2004). A Comparison of Multiple-Choice and Open-Ended Response Formats for the Assessment of Listening Proficiency in English. *Foreign Language Annals*. 37 (4). p. 544-555.
- Chiba, R., Matsuura, H. and Yamamoto, A. (1995). Japanese Attitudes Toward English Accents. *World Englishes*. [Online] ResearchGate 14 (1). p. 77-86. Available from: http://www.researchgate.net/publication/227951660_Japanese_attitudes_toward_English_accents. [Accessed: 5th October 2015].
- Chien, S. (2014). *Varieties of English: Taiwanese Attitudes and Perceptions*. Newcastle and Northumbria Working Papers in Linguistics. [Online] Newcastle and Northumbria Working Papers in Linguistics 20. Newcastle: Newcastle University. Available from: <http://www.ncl.ac.uk/linguistics/assets/documents/SHOUCHUNCHIEN.pdf>. [Accessed: 7th October 2015].
- Clarke, C. M. and Garrett, M. F. (2004). Rapid adaptation to foreign accented English. *Journal of the Acoustical Society of America*. 116 (6). p. 3647-3658.
- Dahl, T. I. & Ludvigsen, S. (2014). How I see what you're saying: The role of gestures in native and foreign language listening comprehension. *Modern Language Journal*. 98 (3). p. 813-833.
- Dalton-Puffer, C., Kaltenboeck, G. and Smit, U. (1997). Learner Attitudes and L2 Pronunciation in Austria. *World Englishes*. [Online] ResearchGate 16 (1). p. 115-128. Available from: http://www.researchgate.net/profile/Ute_Smit/publication/227906705_Learner_Attitudes_and_L2_Pronunciation_in_Austria/links/545bb23a0cf249070a7a7b69.pdf. [Accessed: 5th October 2015].
- Derwing, T. M. and Munro, M. J. (1997). Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition*. 19 (1). p. 1-16.

- Eisenstein, M. R. and Verdi, G. (1985). The intelligibility of social dialects for working-class adult learners of English. *Language Learning*. 35 (2). p. 287-298.
- Floccia, C. et al. (2006). Does a Regional Accent Perturb Speech Processing? *Journal of Experimental Psychology. Human Perception & Performance*. 32 (5). p. 1276-1293.
- Grondelaers, S. and Van Hout, R. (2010). 'Do Speech Evaluation Scales in a Speaker Evaluation Experiment Trigger Conscious or Unconscious Attitudes?' Selected Papers from NWA V 38. [Online] University of Pennsylvania Working Papers in Linguistics 16 (2). Pennsylvania: University of Pennsylvania. Available from: <http://repository.upenn.edu/cgi/viewcontent.cgi?article=1149&context=pwpl>. [Accessed: 24th November 2015].
- Gardner, R. C. (1982). Language attitudes and language learning. In: RYAN, E. B. and GILES, H. (eds.) *Attitudes towards Language Variation. Social and Applied Contexts*. London: Edward Arnold.
- Gass, S. and Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*. Deep Blue 34 (1). p. 65-89. [Online] Available from: <http://deepblue.lib.umich.edu/bitstream/handle/2027.42/98153/j.1467-1770.1984.tb00996.x.pdf?sequence=1>. [Accessed: 19th August 2015].
- Giles, H. 1970. Evaluative reactions to accents. *Educational Review*. [Online] Taylor & Francis Online 22 (3). p. 211-227. Available from: <http://www.tandfonline.com/doi/abs/10.1080/0013191700220301#aHR0cDovL3d3dy50YW5kZm9ubGluZS5jb20vZG9pL3BkZi8xMC4xMDgwLzAwMTMxOTE3MDAyMjAzMDFAQEAw>. [Accessed: 9th October 2015].
- Guichon, N. & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. *System*. 36 (1). p. 85-93.
- Hakala, H. (2007). "Almost as annoying as the Yank; better accent though." - *Attitudes and Conceptions of Finnish Students toward Accents of English*. MA Thesis. University of Helsinki, Department of English.

- Iivanainen, T. (2013). *Aksentit Open Road Course 8 Oral Skills -oppikirjan äänitteillä – valmistavuus haastavien aksenttien ymmärtämiseen*. Pedagogical studies thesis. University of Helsinki, Department of Teacher Education.
- Kennedy, S. and Trofimovich, P. (2008). Intelligibility, Comprehensibility and Accentedness of L2 Speech: The Role of Listener Experience and Semantic Context. *Canadian Modern Language Review*. 64 (3). p. 459-489.
- Kon, C. K. (2002). 'The influence on outcomes of ESL students' performance strategies on a CALL listening comprehension activity.' Retrospective Theses and Dissertations, Iowa State University. Digital Repository (at) Iowa State University. [Online] Available at: <http://lib.dr.iastate.edu/rtd/16109>. [Accessed: 13th October 2015].
- Kujala, O. and Byholm, M. (2014). 'Kiitos vaan improbaturista, Pierre – ranskalaisaksentti ärsytti englannin ylioppilaskokeessa.' *Helsingin Sanomat*, 22nd November. [Online] Available from: <http://www.hs.fi/kaupunki/a1416551716549>. [Accessed: 10th September 2015].
- Ladegaard, H. J. (1998). National Stereotypes and Language attitudes: The perception of British, American and Australian language and culture in Denmark. *Language and Communication*. 18 (4). p. 251-274.
- Liang, D. (2013). 'The Effects of English Audio-Visual Materials on Listening Comprehension from the Perspective of Schema Theory.' International Conference on Education Technology and Information System, Guangzhou University, Guangzhou July 2013. [Online] Atlantis Press. Available at: <http://www.atlantis-press.com/php/pub.php?publication=icetis13&frame=http%3A//www.atlantis-press.com/php/paper-details.php%3Fid%3D7983> [Accessed: 26th October 2015].

- Lopez-Soto, T. and Barrera-Pardo, P. D. (2007). 'Perception of Accent by L2 Students of English: Subjective Preference vs. Objective Intelligibility.' ICPHS XVI, Saarbrücken, 6th August to 10th August 2007. [Online] University of Seville, English Language Department. Available from: <http://www.icphs2007.de/conference/Papers/1625/1625.pdf>. [Accessed: 5th October 2015].
- Louhiala-Salminen, L. and Kankaanranta, A. (2012). Language as an issue in international internal communication: English or local language? If English, what English? *Public Relations Review*. 38 (2). p. 262-269.
- Major, R. C. et al. (2005). Testing the Effects of Regional, Ethnic, and International Dialects of English on Listening Comprehension. *Language Learning*. 55 (1). p. 37-69.
- Matsuura, H., Chiba, R. and Yamamoto, A. (1994). Japanese college students' attitudes toward non-native varieties of English. In: GRADDOL, D. and SWANN, J. (eds.). *Evaluating Language: Papers From the Annual Meeting of the British Association for Applied Linguistics Held at the University of Essex, September 1992*. Clevedon: Multilingual Matters.
- Mayer, R. E. (2001). *Multimedia learning*. 9th print 2007. Cambridge: Cambridge University Press.
- Moinzadeh, A., Rezaei, O. and Dezhara, S. (2012) The Effect of Non-native Accent on Iranian EFL Learners' Listening Comprehension, Focusing on Persian Accent of English. *Journal of Language Teaching and Research*. 3 (5). p. 967-972.
- Opetushallitus. (2015). Lukion opetussuunnitelman perusteet 2015. [Online] Available from: http://www.oph.fi/download/172124_lukion_opetus-suunnitelman_perusteet_2015.pdf. [Accessed: 29th March 2015].
- Paas, F. and Kester, L. (2006). Learners and information characteristics in the design of powerful learning environments. *Applied Cognitive Psychology*. 20 (3). p. 281-285.

- Pihko, M. K. 1994. *Cross-linguistic intelligibility of native and non-native L2 speech varieties*. Jyväskylä: Jyväskylän yliopisto, Korkeakoulujen kielikeskus.
- Pihko, M. K. (1995). *Tavoitteena joustava vieraan kielen puheen ymmärtäminen: Erilaisten englannin puhevarianttien ymmärrettävyys suomenkielisille lukiotason englannin oppijoille*. Jyväskylä: Jyväskylän yliopistopaino.
- Pihko, M. K. (1997). *"His English Sounded Strange": The Intelligibility of Native and Non-Native English Pronunciation to Finnish Learners of English*. PhD Thesis. University of Jyväskylä, Centre for Applied Language Studies.
- Ranta, E. (2010). English in the real world vs. English at school: Finnish English teachers' and students' views. *International Journal of Applied Linguistics*. 20 (2). p. 156–177.
- Rashid, B. I. G. (2011). Foreign Language Accents and EFL Learners' Attitudes. *Journal of Basrah Researches (Humanities Series)*. [Online] Iraqi Academic Scientific Journals 36 (4). p. 58-80. Available from: <http://www.iasj.net/iasj?func=fulltext&aId=58649>. [Accessed: 8th October 2015].
- Rindal, U. (2014). Questioning English standards: Learner attitudes and L2 choices in Norway. *Multilingua*. 33 (3). p. 313-334.
- Ryan, E. B., Giles, H. and SEBASTIAN, R. J. (1982). An integrative perspective for the study of attitudes toward language variation. In: RYAN, E. B. and GILES, H. (eds.). *Attitudes towards Language Variation. Social and Applied Contexts*. London: Edward Arnold.
- Secules, T., Herron, C. and Tomasello, M. (1992). The Effect of Video Context on Foreign Language Learning. *Modern Language Journal*. 76 (4). p. 480-490.
- Smith, L. E. and Bisazza, J. A. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*. 32 (2), p. 259-269.

- Sueyoshi, A. and Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*. 55 (4). p. 661-699.
- Tardieu, H. and Gyselinck, V. (2003). Working memory constraints in the Integration and Comprehension of Information in a Multimedia Context. In: VAN OOSTENDORP, H. (ed.). *Cognition in a Digital World*. London: Lawrence Erlbaum Associates, Publishers.
- Wagner, E. (2010). The effect of the use of video texts on ESL listening test-taker performance. *Language Testing*. 27 (4). p. 493-513.
- Wells, J. C. (1982). *Accents of English I: An Introduction*. Cambridge: Cambridge University Press.
- Weyers, J. R. (1999). The effect of authentic video on communicative competence. *Modern Language Journal*. 83 (3). p. 339-349.
- Wilcox, G. K. (1978) The Effect of Accent on Listening Comprehension: A Singapore Study. *English Language Teaching Journal*. 32 (2). p. 118–127.
- Yanagawa, K. & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple-choice listening comprehension test. *System*. 36 (1). p. 107-122.
- Yi'an, W. M. (1998). What do tests of listening comprehension test? -A retrospective study of EFL test-takers performing a multiple-choice task. *Language Testing*. 15 (1). p. 21-44.

List of the YouTube videos used in the MC test

Hugh Lupton (2013). *Hugh Lupton "Who takes the most Pleasure from the Act of Love?"*. Available from: <https://www.youtube.com/watch?v=lfVzd9IUQVk>. [Accessed 16th November 2015].

SVD Diversity Videos (2013). [106] *My Long Hair*. Available from: <https://www.youtube.com/watch?v=ugGWiQiFsGA>. [Accessed 6th November 2015].

University of Manitoba (2010). *Ignatius Mabasa tells story of the Great Hunter*. Available from: <https://www.youtube.com/watch?v=JqKmpDbkqTU&list=PL3F103BF5517F7631&index=35>. [Accessed 30th October 2015].

WalkingNarrative (2010). *Eddie Lenihan-The Black Dog.mov*. Available from: <https://www.youtube.com/watch?v=dpXnIs57678>. [Accessed 30th October 2015].

Appendix 1. Transcriptions of the texts used in the MC task

The breaks taken during the MC task are marked by three stars (***)

Text 1 (RP): Who takes most pleasure from the act of love? By Hugh Lupton

Zeus and his wife Hera were making love in their golden bed in their palace on the high slopes of Mount Olympus. Zeus the cloud compellor and Hera the queen of heaven. Zeus taking his stormy pleasure of Hera and Hera taking her earthy cunning pleasure of Zeus. And when they'd finished, Zeus rolled over and he fell asleep as he always did.

And Hera shook his shoulder and said: "Zeus, can we talk for once, you and me?"

And Zeus grunted and they sat up side by side on the white pillows. And they talked about this and they talked about that and the conversation came circling round to the question: who takes the most pleasure from the act of love, a man or a woman?

And Zeus said there's only one person who will know the answer for that question and that is Tiresias. He alone has lived both as male and as female. Because there was once a man whose name was Tiresias. And he was walking through the skirts of a forest and he saw two snakes entwined in the act of love. And he thought: here's a chance to kill two serpents with one blow. He reached, he picked up a stick. He lifted it above his shoulder and he brought it down with a thwack onto the backs of the two snakes but they slithered away unharmed.

And as for Tiresias he felt suddenly strangely changed as though all the terminals of pleasure and pain in his body had mysteriously shifted position. And he looked down and he had become a woman. And not only with the body of a woman but with the mind and the heart and the soul of a woman. And for seven years she lived and she loved as a woman. And then seven years later she was walking along the edge of the same forest and she saw the same two snakes entwined in the act of love. And she stood and she watched them for a while. And then out of curiosity she picked up a stick and she brought it down with a thwack onto the backs of the snakes. And he was as male a member of the masculine kind as he had ever been.

And it was on account of that story that Zeus and Hera called to Hermes, the messenger of the gods, and they told him to go and fetch Tiresias. And Hermes strapped on his wing-sandals and swift as thought, he flew down to the world and in a moment he was back with Tiresias and Tiresias stood at the foot of the golden bed shaking and quaking with terror. And Zeus looked him up and down.

“Tell us the truth without beating about the bush: who takes the most pleasure from the act of love, a man or a woman?”

And Tiresias answered as plainly as he knew how: “On a scale of one to ten, a woman regularly reaches nine but a man rarely rises much above five.”

And Hera began to tremble with fury. “How dare you? How dare you reveal the innermost secret of my sex? For all sundry to hear.”

And she reached down to the foot of the bed and she grabbed Tiresias by the hair and with her fingers she gouged out his eyes. And Zeus looked at Tiresias with the blood streaming down his cheeks and he felt pity stirring in his heart. But what one god or goddess has done, cannot be undone by another. And so Zeus reached down to the foot of the bed and with the tip of his finger he touched Tiresias on the forehead. He opened up his inner eye. He gave him the gift of second sight.

He said: “There, the secrets of the future, they are yours now.”

And from that day onwards, Tiresias became the greatest seer, the greatest prophet, ever to have walked the face of the earth.

Text 2 (ZIM): The great hunter by Ignatius Mabasa

I will tell you a story from way back. Umh, when people were still great hunters and there were a lot of animals that – actually, people lived on hunting. Hunting was a profession. It so happened that, a great hunter visited a medicine man and he said to the medicine man:

“Medicine man, look, I know I am the greatest hunter but I want to be even greater. I want to become a living legend.”

But the medicine man looked at the hunter and he said: “I tell you what, you are okay. Don’t ask for more. I think you are great.”

But the medicine man, you know, his words were not well received by the hunter because the hunter said: “The medicine man does not understand what I want. I want more.”

So, he left the house of the medicine man and happy, went to his home. That night he had a dream and in the dream he heard a voice say:

“If you really want to be great, greater than what you are, you can actually go and do something that will make you great. And this something, go down to the river, kill a hippo, or get any hippo, get his tongue. When you have the tongue of the hippo, cook it, eat it and see what happens next. You will be the greatest.”

Now, the hunter, early in the following day, woke up and he said: “I am going to follow my dream. I’m going to do exactly what I heard in the dream.”

So, he took his spear and he walked to the river. He went to this huge river that was in the village and there he saw early in the morning crocodiles basking in the sun and hippos trying to get something to eat early in the morning. As he walked he looked at some of the hippos. They really looked intimidating and formidable and he’s like:

“Hmm... We’ll see.”

But he was a lucky man because soon he found an old hippo lying in the mud, dying and he said: “I may as well finish this dying hippo off.”

So, he threw his spear and speared this hippo and it died. Now he said:

“I will have what I wanted which is the tongue.”

So, he walked over, tried to open the mouth of the hippo, but you know, the mouth of the hippo is so huge it’s more like a cave and he struggled but the mouth wouldn’t open. So, he summoned all the energies and powers that he had, pushed hard and slowly the mouth opened a little bit.

And he said: “I must find something to prop and keep this mouth, you know, open.”

So, he took a small rock, placed it between the teeth of the hippo and the mouth remained open a little bit.

Then he said: “I must still open this mouth wider, so that I can actually walk in and cut the tongue, but how am I going to do that?”

He said to himself: “Alright, I’ve got my spear.”

So, he used the spear and pushed the mouth wide open, pushed so hard that finally when the mouth was fully open, he propped the mouth open with his spear and he said:

“This is it. Now we’re talking business.”

He took the knife from his pocket and he crawled into the mouth of the hippo. And he really wanted to get a big chunk of the tongue and he really pushed himself further. As he kicked to push himself further to the back of the mouth so that he could get a good grip of the tongue, he kicked the spear. And because it was propping the mouth open, the mouth just shut like that with him inside. That became the end of the great hunter and his quest for greatness. Thank you.

Text 3 (JPN): My long hair by Motoko Dworkin

Hi. My name is Motoko. As a teenager growing up in Osaka, Japan, I was not pretty or popular but my hair was. Yes, I used to have this long, shiny, silky black hair straight down to my waist. So much of it I would obsessively brush it to an arresting shine. During the school day the rules dictated that I had to keep it in a single long braid. As soon as the school let out, I would untie my braid and shake it loose into a shimmering cascade. What a glorious feeling.

I was a good studious student. By the time I was in the tenth grade, I was attending what we call ‘yobiko’, a cram school after the regular high school three nights a week for extra math and English lessons to prepare myself for the college entrance exams. On those days, I did not get to let my hair down until much later because those classes did not finish until nine o’clock at night. Then I would take the commuter train home, get home about ten o’clock, eat dinner and do my homework.

One night I was on my way home. As usual, the commuter train was jam-packed with businessmen and laborers, some drunken and boisterous, others tired and sullen. A few of them leered at me, a girl in a school uniform with long hair in a single braid. I sat with my knees together with a heavy book bag on my lap. By the time I finally got off the train, the crowd had thinned a little. I walked toward the bicycle lot at the back of the station away from the blaring music and the neon signs of karaoke bars and pachinko parlors. I found my bike and dropped my heavy bag in a red wire basket attached in front of it. With relief I untied my braid and swung my head letting the spring breeze cool down my scalp.

“Nice hair!” A man’s guttural voice from right behind startled me. I spun around into reeky fumes of hot drunken breath. A stranger sneering too close. Shallow cheeks and a stubby chin, a dark green shirt. The next thing, he was grabbing my waist, pulling me hard against him. No, I did not scream. I was too stunned to even make a sound. The whole thing seemed somehow not so real, like a scene from a silent movie. I struggled to free myself and the man suddenly let loose. I staggered back and bumped into my bike. The bike fell and I fell on top of it, scraping my leg against the pedal. My books were scattered everywhere. Then the man suddenly started to laugh hysterically as if he had never seen anything so funny.

I rided my bike and without looking back, pedaled as fast as I could. When I finally reached my house, I realized that I had not breathed the whole time. I got off my bike and bent down to waist as if I had just sprinted a mile. My heart was beating so fast in my head I could not hear anything else. No, the man did not follow me. I only had two bruises and a long scratch on my leg. My blouse had come untucked, so I tucked it in. I was okay, nothing happened. I opened the door. The glaring fluorescent light and the smell of the dinner and the loud noise from the TV and the living room overwhelmed me.

Text 4 (IRE): The black dog by Eddie Lenihan

Umh, what I have been interested in for quite a number of years is the fairies. So, an old man said to me once, and he was a reliable old man, a man who wasn’t giving all of the nonsense. I think I have a hundred and fifty hours of him on tape. He was an encyclopedia of knowledge about local history and about this kind of thing and...

And he said to me once, he said: “Look, the fairies are like us.” He said: “The person sitting next to you could be one of them and you wouldn’t know it.”

And that’s a frightening thought. They can take any shape they like and one of their favorite shapes according to Irish tradition is the black dog.

He said that’s where on that road bridge the black dog used to be seen. He was a monster. He was as big as a horse. And he’d plunk himself there on his backside on the bridge, on his backside out for the wall on either side.

And fellas coming back at night from the pub or from card play or from visiting their neighbors crossed their souls as they’d come for this monster. What were you going to do? Ask him: “Hello, hello little doggy can I pass you by?” They were frightened out of their wits for him. And so they would have to go two or three miles out of their way by other roads or, or more dangerously still, across the river upstream or downstream which was okay in the summertime when the river was low. But in the wintertime when the river was flooded, you could be drowned. Now, there was a plank there, across. But I mean, crossing by a plank on a flooded night – dangerous, dangerous.

So, eventually they said: “Look, what we do about this? Who could help us?”

And they decided that the obvious man was the priest, the parish priest. So, they went to him. Huh, and you know the reception by God. A lecture.

“Why in the name of God almighty don’t you stay out of the pubs and you wouldn’t be seeing black dogs. Nonsense, Christ Jesus! Will you go home and have a small bit of sense.”

And they still had to travel the road at night. So, they were still having to go around and eventually they sent their wives to the priest.

Now, now, that was different. Because in those days you might laugh at a man but women weren’t going to the pubs in Ireland. They might be having their little drop at home but they weren’t going to the pubs. And when the women, the wives, started coming to the presbytery to the priest’s house, he began to take notice. And he began to think maybe there is something in this.

“So, be darned,” he said. “Look, it will do no harm to check it out.”

So, that very night when the women were gone about their business, he went, armed with holy water and his book, prayer book, I presume, came to the bridge, but nothing.

But he was a careful kind of a man and he came the second night. Nothing. Third night, nothing.

But the fourth day was a Sunday. So, he said he'd put an end to this once and for all. So, God almighty, at mass that Sunday, he preached them a sermon of the... out of calling them all kinds of superstitious peasants and blah blah blah. All that kinds of stuff. And of course they went out of there feeling like that height.

But yet, aye, the dog was still there and the women were still coming to the priest night after night after night after night. And he went back, he went back. Fourth night, no. Fifth night, just to keep the women quiet at this stage. Sixth night, nothing.

On the seventh night, didn't he meet the dog. Lucky for himself, he had his holy water and his book. And the man that was telling me the story said that the neighbors all around they heard the yelping. They heard the screeching, screeching ten times as loud as any ordinary dog. And of course, nobody went out to investigate. They figured out it must be the priest had met the animal.

And it was only in the following morning, a man going out to collect his cows to bring them in to milk them, down near where the stream went into the sea – I won't even call it an estuary because the stream is only about four miles long, small. There he found the dog, inside in the water, dead, with its legs up in the air. And of course the news spread. People came from ten miles around to see this creature. What in God's name was it? Oh yeah, a dog, but the size of a horse.

They knew what it was of course. They knew that this was no ordinary dog and as the man said:

“My father was there that morning. He saw it with his own two eyes. He told me the story.”

And after a few more days, high tide came in and washed the dog, or whatever it was, out and it was seen no more.

But he says: “Anybody who will tell you that these things don't exist doesn't know what he's talking about. Or that it means is he's never met them. But anybody who has, I tell you, they have a different idea, the fairies and this little 'woo'.”

Appendix 2. The MC test in Finnish

The 1st grade MC test is presented first after which the 3rd grade MC test is presented.

Tutkimus toteutetaan anonymisti.

Ikä: ___ **Sukupuoli:** mies nainen **Ryhmä:** 1 2

Ohjeet: Ympyröi oikea vaihtoehto. Tekstit kuunnellaan vain kerran. Tekstissä pidetään 30 sekunnin tauko kahden kysymyksen välein. Tekstin loputtua pidetään kahden minuutin tauko, jonka aikana vastataan kysymyksiin kertojasta ja valmistaudutaan seuraavan tekstin ensimmäisiin kysymyksiin. Kiitos osallistumisestasi!

Teksti 1

1. Mitä Hera halusi tehdä rakastelun jälkeen?
 - a. Rakastella uudestaan.
 - b. Käydä nukkumaan.
 - c. Keskustella.
 - d. Lähteä maan kamaralle.

2. Mihin kysymykseen Zeus ja Hera päätyivät?
 - a. Onko mies vai nainen vahvempi?
 - b. Kenellä on kovimmat paineet?
 - c. Onko ympyrä oikeasti pyöreä?
 - d. Kuka nauttii eniten rakastelusta?

3. Miksi Tiresias tietäisi vastauksen kysymykseen?
 - a. Koska hän oli tutkinut käärmeiden parittelua.
 - b. Koska hän tunsu miehet ja naiset parhaiten.
 - c. Koska hän oli elänyt sekä miehenä että naisena.
 - d. Koska hän oli kaikkein viisain ihminen.

4. Mitä Tiresias teki kohdatessaan käärmeet?
 - a. Löi niitä löytämällään kepillä.
 - b. Alkoi pohtimaan rakkauden luonnetta.
 - c. Käveli pois päin, jotteivat ne purisi.
 - d. Poimi ne käteensä ja heitti maahan.

5. Mitä käärmeet tekivät, kun Tiresias kohtasi ne uudestaan?
 - a. Rakastelivat kietoutuneina toisiinsa.
 - b. Makasivat paikoillaan tuijottaen Tiresiasta.
 - c. Syöksyivät eteenpäin ja purivat Tiresiasta.
 - d. Luikertelivat tiehensä.

 6. Millainen Tiresias oli sen jälkeen, kun hän oli kohdannut käärmeet toisen kerran?
 - a. Mietteliäs ja vaitonainen.
 - b. Niin miehekäs kuin hän oli eläessään ollut.
 - c. Vanha, mutta entistä viisaampi.
 - d. Pettynyt elämänsä käännteisiin.
-

7. Millainen oli Tiresiaksen reaktio hänen kohdatessaan jumalat?
 - a. Hän tärisi kauhusta.
 - b. Hän meni sanattomaksi.
 - c. Hän riemuitsi.
 - d. Hänen ajatuksensa alkoivat juosta nopeammin.

 8. Miten Tiresias vastasi jumalten kysymykseen?
 - a. Mies nauttii enemmän.
 - b. Nainen nauttii enemmän.
 - c. Mies ja nainen nauttivat yhtä paljon.
 - d. Hän ei tiennytkään vastausta.
-

9. Mitä Hera teki Tiresiakselle?
 - a. Katkoi häneltä sormet.
 - b. Repi häneltä hiukset päästä.
 - c. Kaivoi häneltä silmät päästä.
 - d. Mursi hänen jalkansa.

10. Mitä Zeus teki Tiresiakselle?
 - a. Pyyhki veren hänestä samalla parantaen hänet.
 - b. Lähetti hänet tulevaisuuteen, jotta hän voisi löytää parannuskeinon.
 - c. Lähetti hänet takaisin maan kamaralle.
 - d. Avasi hänen sisäisen silmänsä koskettamalla hänen otsaansa.

Teksti 2

1. Millaisesta ajasta kertojan tarina kertoo?
 - a. Ajasta, jolloin ihmiset pelkäsivät eläimiä.
 - b. Ajasta, jolloin eläimet osasivat puhua.
 - c. Ajasta, jolloin metsästys oli vielä ihmisten elinkeino.
 - d. Ajasta, jolloin eläimet metsästäivät ihmisiä.

2. Mitä mieltä suuri metsästäjä oli poppamiehen neuvoista?
 - a. Neuvot olivat hyviä, mutta metsästäjä halusi vielä lisää.
 - b. Hän päätti noudattaa neuvoja.
 - c. Poppamiehen tarjoama lääke ei riittänyt hänen mielestään.
 - d. Ne eivät kelvanneet hänelle.

3. Mitä metsästäjän piti unessa kuuluneen ääneen mukaan tehdä, jotta hänestä tulisi mahtava?
 - a. Valmistaa ateria virtahevolle.
 - b. Hänen piti tehdä sankariteko.
 - c. Hänen pitäisi odottaa, ja hänen tilaisuutensa koittaisi.
 - d. Syödä virtahevon kieli.

4. Milloin metsästäjä heräsi?
 - a. Aikaisin seuraavana päivänä.
 - b. Keskellä yötä, heti unen jälkeen.
 - c. Vasta päivien päästä.
 - d. Seuraavan päivän iltana.

5. Millaisia virtahevot olivat?
 - a. Säyseitä.
 - b. Pelottavia.
 - c. Ihmisystävällisiä.
 - d. Vanhoja.

6. Miksi metsästäjä oli onnekas?
 - a. Hän löysi mutaa, jolla naamioi itsensä.
 - b. Koska hän löysi virtahevot niin helposti.
 - c. Hän löysi kuolevan virtahevon.
 - d. Hänellä ei ollut kiire lopettaa etsintää.

7. Miksi virtahevon suu oli vaikeaa avata?
 - a. Se oli suuri kuin luola.
 - b. Suu oli juuttunut kiinni.
 - c. Virtahevon leuoista ei saanut kunnon otetta.
 - d. Se oli täynnä hampaita.

 8. Miten metsästäjä sai virtahevon suun pysymään raollaan?
 - a. Hampaat asettuivat vastakkain pitäen suuta auki.
 - b. Hän asetti kiven hampaiden väliin.
 - c. Suu pysyi raollaan kuolemanjäykkyyden ansiosta.
 - d. Hänen täytyi käyttää kaikki voimansa, jotta hän sai pidettyä suun raollaan.
-

9. Miten metsästäjä sai virtahevon suun pysymään kokonaan auki?
 - a. Suu oli niin jäykkä, että se pysyi auki.
 - b. Hän pystyi pitämään suun auki selällään.
 - c. Hän pönkitti keihäänsä leukojen väliin.
 - d. Hän asetti veitsensä hampaiden väliin.

10. Miten metsästäjä pääsi virtahevon suuhun?
 - a. Hän veti itseään syvemmälle kieltä pitkin.
 - b. Hän käveli sisään.
 - c. Hän potki esteet tieltään.
 - d. Hän ryömi sisään.

Teksti 3

1. Mitä Motoko teki teininä tukalleen?
 - a. Leikkasi sen vyötärölle ulottuvaksi.
 - b. Kokeili erilaisia kampauksia.
 - c. Pesi sen usein, jotta se kiiltäisi.
 - d. Harjasi sitä pakkomielteisesti.

2. Mitä Motoko teki koulupäivien jälkeen?
 - a. Letitti hiuksensa yhdelle pitkälle letille.
 - b. Hän avasi palmikkonsa ja ravisti hiukset avoimiksi.
 - c. Kampasi hiuksensa, jotta ne näyttäisivät paremmilta.
 - d. Pureskeli tukkaansa, koska siitä tuli hänelle hyvä olo.

3. Mitä Motoko opiskeli yobikossa eli ylimääräisessä koulussa?
 - a. Kemiaa ja englantia.
 - b. Äidinkieltä ja englantia.
 - c. Matematiikkaa ja englantia.
 - d. Saksaa ja englantia.

4. Mihin aikaan yobiko loppui?
 - a. Koulusta pääsi, kun tehtävät oli tehty.
 - b. Yhdeksältä illalla.
 - c. Kymmeneltä illalla.
 - d. Loppumisaika riippui päivästä.

5. Mitä oli tapahtunut siihen mennessä, kun Motoko nousi pois junasta?
 - a. Väenpaljous oli harventunut.
 - b. Juna oli pikkuhiljaa pakkautunut täpötäyteen ihmisiä.
 - c. Raskas kirjalaukku alkoi painaa Motokon jalkoja.
 - d. Motoko oli vaihtanut koulupukunsa normaaleihin vaatteisiin.

6. Mihin Motoko laittoi kirjalaukkunsa?
 - a. Hän unohti laukun junaan.
 - b. Karaokebaarin viereiseen säilytyslokeroon.
 - c. Väsyneenä hän tiputti sen vahingossa roskakoriin.
 - d. Polkupyöräänsä kiinnitettyyn koriin.

7. Miksei Motoko huutanut jonkun tarttuessa häneen?
 - a. Hän tiesi, ettei kukaan kuulisi häntä siinä melussa.
 - b. Mies peitti hänen suunsa kädellään.
 - c. Hän oli liian ällistynyt huutaakseen.
 - d. Mies kielsi häntä päästävästä ääntäkään.

 8. Miten Motokon jalka naarmuttui?
 - a. Mies raapaisi häntä heidän kamppaillessaan.
 - b. Jalka raapaisi maata hänen kaatuessaan.
 - c. Se osui hänen polkupyöränsä polkimeen.
 - d. Hänen kirjansa tippuivat jalan päälle.
-
9. Mitä Motoko tajusi, kun hän oli päässyt kotiovelleen asti.
 - a. Hänen vyötäröään oli alkanut särkeä.
 - b. Hän ei ollut hengittänyt koko aikana.
 - c. Hän oli juossut ainakin mailin.
 - d. Hän oli menettänyt kuulonsa.

 10. Mitä Motokon puserolle oli käynyt?
 - a. Se oli auennut.
 - b. Se oli revennyt.
 - c. Se oli liian peitossa.
 - d. Puserolle ei käynyt mitään.

Teksti 4

1. Mitä mies oli sanonut keijuista?
 - a. Vieressäsi istuva henkilö voisi olla yksi heistä.
 - b. Keijut pitävät ihmisistä.
 - c. Niillä on usein lemmikkinä mustia koiria.
 - d. Keijut tietävät ihmisistä kaiken.

2. Mitä musta koira teki?
 - a. Jahtasi pubista tulevia ihmisiä.
 - b. Kirosi ihmisten sielut.
 - c. Tukki sillan takamuksellaan.
 - d. Puhui ihmisille öisin.

3. Mitä ihmiset tekivät välttyäkseen kohtaamasta mustaa koiraa?
 - a. Kulkivat kiertoteitä tai ylittivät joen eri kohdasta.
 - b. Liikkuivat ulkona vain kesäisin ja pysyttelivät sisällä talvisin.
 - c. Odottivat joen tulvimista, koska silloin koiraa ei näkynyt.
 - d. Pelottelivat koiraa huutelemalla.

4. Mikä oli papin reaktio, kun häneltä pyydettiin apua?
 - a. Pappikin pelkäsi koiraa niin, ettei halunnut auttaa.
 - b. Hän saarnasi pubeissa notkumisesta.
 - c. Pappi rukoili apua Jumalalta.
 - d. Hän yritti lohduttaa ihmisiä.

5. Miksi naiset lähetettiin papin puheille?
 - a. Koska papin tiedettiin ihastuneen erääseen kauniiseen naiseen.
 - b. Heidän piti päästä välillä pois kotoa.
 - c. Naiset olivat hartaampia uskovaisia.
 - d. Koska siihen aikaan naiset eivät käyneet pubeissa.

6. Mitä pappi otti mukaansa etsiessään mustaa koiraa?
 - a. Hän ei ottanut mukaansa mitään.
 - b. Aseen ja ammuksia.
 - c. Yhden kylän naisista.
 - d. Vihkivettä ja rukouskirjan.

7. Mitä pappi teki neljäntenä päivänä?
 - a. Etsi mustaa koiraa koko päivän ja yön.
 - b. Lepäsi, koska neljäs päivä oli sunnuntai.
 - c. Hän piti saarnan, jossa haukkui ihmisiä taikauskaisiksi.
 - d. Rukoili apua tilanteeseen koko päivän.

 8. Mistä ihmiset tiesivät papin kohdanneen mustan koiran?
 - a. He menivät ulos tutkimaan, mitä siellä oikein tapahtui.
 - b. He kuuluivat ulvontaa, joka oli paljon kovempaa kuin normaalin koiran ulvonta.
 - c. He tiesivät, että koira näyttäytyy vähintään seitsemän päivän välein.
 - d. He seurasivat pappia salaa ja näkivät kohtaamisen.
-
9. Mitä musta koira teki, kun se kohdattiin seuraavana päivänä?
 - a. Koira makasi kuolleena joessa jalat kohti taivasta.
 - b. Pureskeli tappamansa lehmän ruumista.
 - c. Jahtasi ihmisten hevosia laitumella.
 - d. Juoksi tiehensä eikä sitä enää koskaan nähty.

 10. Mitä kertoja sanoo ihmisistä, jotka sanovat, ettei mustia hirviökoiria ole olemassa?
 - a. He ovat oikeassa, koska vain hullut ovat taikauskaisia.
 - b. He eivät ole koskaan nähneet sellaista koiraa, eivätkä tiedä mistä puhuvat.
 - c. Heidän pitäisi osata erottaa sadut tositarinoista.
 - d. Heidän pitäisi uskoa, mitä vanhemmat ihmiset heille kertovat.

3rd grade questions**Teksti 1**

1. Mitä Hera halusi tehdä rakastelun jälkeen?
 - a. Rakastella uudestaan.
 - b. Käydä nukkumaan.
 - c. Keskustella.
 - d. Lähteä maan kamaralle.

2. Mihin kysymykseen Zeus ja Hera päätyivät?
 - a. Rakastuuko mies vai nainen helpommin?
 - b. Kenelle rakastelu on tärkeintä?
 - c. Onko miehen vai naisen rakkaus aidompaa?
 - d. Kuka nauttii eniten rakastelusta?

3. Miksi Tiresias tietäisi vastauksen kysymykseen?
 - a. Koska hän oli tutkinut käärmeiden parittelua.
 - b. Koska hän tunsu miehet ja naiset parhaiten.
 - c. Koska hän oli elänyt sekä miehenä että naisena.
 - d. Koska hän oli kaikkein viisain ihminen.

4. Mitä Tiresias teki kohdatessaan käärmeet?
 - a. Löi niitä löytämällään kepillä.
 - b. Alkoi pohtimaan rakkauden luonnetta.
 - c. Käveli pois päin, jotteivat ne purisi.
 - d. Poimi ne käteensä ja heitti maahan.

5. Mitä käärmeet tekivät, kun Tiresias kohtasi ne uudestaan?
 - a. Rakastelivat kietoutuneina toisiinsa.
 - b. Makasivat paikoillaan tuijottaen Tiresiasta.
 - c. Syöksyivät eteenpäin ja purivat Tiresiasta.
 - d. Luikertelivat tiehensä.

6. Millainen Tiresias oli sen jälkeen, kun hän oli kohdannut käärmeet toisen kerran?
 - a. Mietteliäs ja vaitonainen.
 - b. Niin miehekäs kuin hän oli eläessään ollut.
 - c. Vanha, mutta entistä viisaampi.
 - d. Pettynyt elämänsä käännteisiin.

7. Millainen oli Tiresiaksen reaktio hänen kohdatessaan jumalat?
 - a. Pelko veti hänet ensin sanattomaksi.
 - b. Hän ei uskaltanut nostaa katsettaan.
 - c. Hän tärisi kauhusta.
 - d. Hän piiloutui kauhuissaan sängyn taakse.

 8. Miten Tiresias vastasi jumalten kysymykseen?
 - a. Hän yritti kierrellä vastauksen asettelussa.
 - b. Niin yksinkertaisesti kuin osasi.
 - c. Hän yritti valehdella miellyttääkseen jumalia.
 - d. Hän ei osannut esittää suoraa vastausta.
-

9. Mitä Hera teki Tiresiakselle?
 - a. Iski häntä silmille sokeuttaen hänet.
 - b. Puhkoi hänen silmänsä sormillaan.
 - c. Kaivoi häneltä silmät päästä paljain käsin.
 - d. Pakotti Tiresiaksen sokeuttamaan itsensä.

10. Mitä Zeus teki Tiresiakselle?
 - a. Paransi hänet sormensa kosketuksella.
 - b. Zeus ei voinut auttaa häntä.
 - c. Antoi hänelle toiset silmät.
 - d. Avasi hänen sisäisen silmänsä.

Teksti 2

1. Mitä sanotaan tarinan miljööstä (aika- ja paikka)?
 - a. Tarina sijoittuu kaukaiseen paikkaan.
 - b. Tarina sijoittuu läheiseen paikkaan.
 - c. Tarina sijoittuu kaukaiseen aikaan.
 - d. Tarinan tapahtumista ei ole kovin kauaa.

2. Mitä mieltä suuri metsästäjä oli poppamiehen neuvoista?
 - a. Neuvot olivat hyviä, mutta metsästäjä halusi vielä lisää.
 - b. Hän päätti noudattaa neuvoja.
 - c. Poppamiehen tarjoama lääke ei riittänyt hänen mielestään.
 - d. Ne eivät kelvanneet hänelle.

3. Mikä tekisi unessa kuuluneen äänen mukaan metsästäjästä mahtavan?
 - a. Virtahevon tappaminen.
 - b. Virtahevon lihan kokkaaminen.
 - c. Virtahevon kielen irti leikkaaminen.
 - d. Virtahevon kielen syöminen.

4. Milloin metsästäjä heräsi?
 - a. Aikaisin seuraavana päivänä.
 - b. Keskellä yötä, heti unen jälkeen.
 - c. Vasta päivien päästä.
 - d. Seuraavan päivän iltana.

5. Millaisia virtahevot olivat?
 - a. Säyseitä.
 - b. Pelottavia.
 - c. Ihmisystävällisiä.
 - d. Vanhoja.

6. Miksi metsästäjä oli onnekas?
 - a. Yksi virtahepo oli juuttunut mutaan.
 - b. Koska hän löysi virtahevot niin helposti.
 - c. Hän löysi kuolevan virtahevon.
 - d. Virtahevot olivat heikentyneet vanhuutensa vuoksi.

7. Miksi virtahevon suu oli vaikeaa avata?
 - a. Se oli liian suuri.
 - b. Metsästäjä ei ollut tarpeeksi vahva.
 - c. Metsästäjä oli liian väsynyt.
 - d. Virtahepo kamppaili vastaan.

 8. Miten metsästäjä sai virtahevon suun pysymään raollaan?
 - a. Hampaat asettuivat vastakkain pitäen suuta auki.
 - b. Hän asetti kiven hampaiden väliin.
 - c. Suu pysyi raollaan kuolemanjähkytyksen ansiosta.
 - d. Hänen täytyi käyttää kaikki voimansa, jotta hän sai pidettyä suun raollaan.
-

9. Miten metsästäjä sai virtahevon suun pysymään kokonaan auki?
 - a. Hän kampsu suun auki keihäällänsä.
 - b. Hän työnsi sen auki niin kovaa kuin jaksoi.
 - c. Hän pönkitti keihänsä leukojen väliin.
 - d. Hän asetti veitsensä hampaiden väliin.

10. Miten metsästäjä pääsi virtahevon suuhun?
 - a. Hän veti itseään syvemmälle kieltä pitkin.
 - b. Hän käveli sisään.
 - c. Hän potki esteet tieltään.
 - d. Hän ryömi sisään.

Teksti 3

1. Mitä Motoko teki teininä tukalleen?
 - a. Leikkasi sen vyötärölle ulottuvaksi.
 - b. Kokeili erilaisia kampauksia.
 - c. Pesi sen usein, jotta se kiiltäisi.
 - d. Harjasi sitä pakkomielteisesti.

2. Mitä Motoko teki koulupäivien jälkeen?
 - a. Letitti hiuksensa yhdelle pitkälle letille.
 - b. Hän avasi palmikkonsa.
 - c. Kiillotti hiuksensa.
 - d. Heilutteli lettiään.

3. Mitä Motoko opiskeli yobikossa eli ylimääräisessä koulussa?
 - a. Kemiaa ja englantia.
 - b. Äidinkieltä ja englantia.
 - c. Matematiikkaa ja englantia.
 - d. Saksaa ja englantia.

4. Mihin aikaan Motoko pääsi kotiin päivinä, joina hän kävi yobikossa?
 - a. Vaihtelevasti sen mukaan, moneltako yobiko loppui.
 - b. Yhdeksältä illalla.
 - c. Kymmeneltä illalla.
 - d. Vaihtelevasti riippuen siitä, kuinka nopeasti hän sai läksynsä tehtyä.

5. Mitä oli tapahtunut siihen mennessä, kun Motoko nousi pois junasta?
 - a. Väenpaljous oli harventunut.
 - b. Juna oli pikkuhiljaa pakkautunut täpötäyteen ihmisiä.
 - c. Raskas kirjalaukku alkoi painaa Motokon jalkoja.
 - d. Motoko oli vaihtanut koulupukunsa normaaleihin vaatteisiin.

6. Mihin Motoko pudotti kirjalaukkunsa?
 - a. Polkupyörien parkkialueelle.
 - b. Rautalangasta tehtyyn säilytyslokeroon.
 - c. Punaiseen roskakoriin.
 - d. Polkupyöränsä koriin.

7. Miksei Motoko huutanut jonkun tarttuessa häneen?
 - a. Tilanne vaikutti niin epätodelliselta.
 - b. Hän yritti, mutta ääntä ei tullut.
 - c. Hän oli liian ällistynyt.
 - d. Hän keskittyi rimpuilemaan.

 8. Miten Motokon jalka naarmuttui?
 - a. Mies raapaisi häntä heidän kamppaillessaan.
 - b. Jalka raapaisi maata hänen kaatuessaan.
 - c. Se osui hänen polkupyöränsä polkimeen.
 - d. Hänen kirjansa tippuivat jalan päälle.
-

9. Miksi Motokon piti kumartua kotiovellaan?
 - a. Hänen vyötäröään oli alkanut särkeä.
 - b. Hän ei ollut hengittänyt paetessaan.
 - c. Hän oli juossut mailin.
 - d. Hänen päätänsä oli alkanut särkeä.

10. Mitä Motokon puserolle oli käynyt?
 - a. Se oli auennut.
 - b. Se oli revennyt.
 - c. Se oli liian peitossa.
 - d. Puserolle ei käynyt mitään.

Teksti 4

1. Millaisia keijut miehen mukaan olivat?
 - a. Ne istuskelevat usein ihmisten lähellä.
 - b. Keijut pitävät ihmisistä.
 - c. Niitä voisi luulla ihmisiksi.
 - d. Keijut tietävät ihmisistä kaiken.

2. Mitä musta koira teki?
 - a. Jahtasi pubista tulevia ihmisiä.
 - b. Kirosi ihmisten sielut.
 - c. Tukki sillan takamuksellaan.
 - d. Puhui ihmisille öisin.

3. Mitä ihmiset tekivät välttyäkseen kohtaamasta mustaa koira?
 - a. He liikkuiivat ulkona vain kesäisin.
 - b. Odottivat joen tulvimista.
 - c. Muuttivat kulkureittiään.
 - d. He liikkuiivat ulkona vain talvisin.

4. Mikä oli papin reaktio, kun häneltä pyydettiin apua?
 - a. Pappikin pelkäsi koiraan niin, ettei halunnut auttaa.
 - b. Hän kutsui ihmiset saarnatilaisuuteen.
 - c. Hän kehotti ihmisiä pysymään sisällä pubeissa.
 - d. Hän saarnasi kapakoissa notkumisesta.

5. Miksi naiset lähetettiin papin puheille?
 - a. Koska papin tiedettiin ihastuneen erääseen kauniiseen naiseen.
 - b. Heidän piti päästä välillä pois kotoa.
 - c. Naiset olivat hartaampia uskovaisia.
 - d. Koska siihen aikaan naiset eivät käyneet pubeissa.

6. Miten pappi aseistautui etsiessään mustaa koira?
 - a. Hän ei ottanut mukaansa mitään.
 - b. Kertoja ei tiennyt, mutta oletti papin olevan aseistautunut.
 - c. Hän otti mukaansa pyhiä esineitä.
 - d. Hänellä oli vihkivettä ja rukouskirja.

7. Mitä pappi teki neljäntenä päivänä?
 - a. Etsi mustaa koiraa koko päivän ja yön.
 - b. Lepäsi, koska neljäs päivä oli sunnuntai.
 - c. Hän piti saarnan, jossa haukkui ihmisiä taikauskoisiksi.
 - d. Rukoili apua tilanteeseen koko päivän.

 8. Mistä ihmiset tiesivät papin kohdanneen mustan koiran?
 - a. He menivät ulos tutkimaan, mitä siellä oikein tapahtui.
 - b. He kuuluivat ulvontaa, joka oli paljon kovempaa kuin normaalin koiran ulvonta.
 - c. He tiesivät, että koira näyttäytyy vähintään seitsemän päivän välein.
 - d. He seurasivat pappia salaa ja näkivät kohtaamisen.
-
9. Miten musta koira kohdattiin seuraavana päivänä?
 - a. Koira löytyi kuolleena joesta.
 - b. Koira leijui ilmassa joen päällä.
 - c. Koira sukelteli joessa.
 - d. Koira oli siirtynyt joen rannalle lähelle merta.

 10. Mitä kertoja sanoo ihmisistä, jotka sanovat, ettei mustia hirviökoiria ole olemassa?
 - a. He ovat oikeassa, koska vain hullut ovat taikauskoisia.
 - b. He eivät ole koskaan nähneet sellaista koiraa, eivätkä tiedä mistä puhuvat.
 - c. Heidän pitäisi osata erottaa sadut tositarinoista.
 - d. Heidän pitäisi uskoa, mitä vanhemmat ihmiset heille kertovat.

Appendix 3. The MC test in English

The 1st grade MC test is presented first after which the 3rd grade MC test is presented.

The research is carried out anonymously.

Age: ___ **Gender:** male female **Group:** 1 2

Instructions: Circle the correct alternative. The texts will be listened to only once. 30-second breaks will always be taken after two questions. After the text has ended, a two-minute break will be taken. You are to answer the questions about the storyteller during the break as well as to prepare yourself for the first questions of the next text. Thank you for your participation!

Text 1

1. What did Hera want to do after making love?
 - a. Make love again.
 - b. Go to sleep.
 - c. Have a conversation.
 - d. Go to the face of the earth.

2. What was the question Zeus and Hera ended up with?
 - a. Who is stronger, man or woman?
 - b. Who has the most pressure?
 - c. Is circle really round?
 - d. Who takes most pleasure from the act of love?

3. Why would Tiresias know the answer to the question?
 - e. Because he had studied the mating of snakes.
 - f. Because he knew men and women the best.
 - g. Because he had lived both as a man and a woman.
 - h. Because he was the wisest human of all.

4. What did Tiresias do when he encountered the snakes?
 - a. He hit them with a stick he found.
 - b. He started to ponder on the nature of love.
 - c. He walked away so that they wouldn't bite him.
 - d. He grabbed them and threw them to the ground.

5. What were the snakes doing when Tiresias encountered them again?
 - a. They were making love entwined to each other.
 - b. They lied still staring at Tiresias.
 - c. They lunged forward biting Tiresias.
 - d. They slithered away.

 6. How was Tiresias like after encountering the snakes for the second time?
 - a. Thoughtful and quiet.
 - b. As manly as he had ever been.
 - c. Old, but wiser than before.
 - d. Disappointed at the twists in his life.
-

7. How was Tiresias' reaction like when he met the gods?
 - a. He trembled with fear.
 - b. He went all speechless.
 - c. He rejoiced.
 - d. His thoughts started to flow faster.

 8. How did Tiresias answer the gods' question?
 - a. Men enjoy more.
 - b. Women enjoy more.
 - c. Men and women enjoy equally.
 - d. He didn't know the answer after all.
-

9. What did Hera do to Tiresias?
 - a. She cut his fingers.
 - b. She tore his hair from his head.
 - c. She gouged out his eyes.
 - d. She broke his legs.

10. What did Zeus do to Tiresias?
 - a. He wiped the blood off him curing him.
 - b. He sent him to the future so that he could find a cure.
 - c. He sent him back to the face of the earth.
 - d. He opened up his inner eye by touching his forehead.

Text 2

1. What kind of time is the storyteller's story about?
 - a. A time when people were afraid of animals.
 - b. A time when animals could speak.
 - c. A time when hunting was still a livelihood for people.
 - d. A time when animals hunted people.

2. What did the great hunter think of the medicine man's advice?
 - a. The advice was good, but the hunter wanted more.
 - b. He decided to follow the advice.
 - c. The medicine the witchdoctor offered wasn't enough in the hunter's opinion.
 - d. The medicine man's advice was not good enough for him.

3. What was the hunter supposed to do according to the voice he heard in the dream?
 - a. Prepare a meal for a hippo.
 - b. He was to perform a heroic deed.
 - c. He should wait and his opportunity would manifest.
 - d. Eat the tongue of a hippo.

4. When did the hunter wake up?
 - a. Early in the following day.
 - b. In the middle of the night, right after the dream.
 - c. Not until days had passed.
 - d. In the evening of the next day.

5. How were the hippos like?
 - a. Docile.
 - b. Scary.
 - c. People-friendly.
 - d. Old.

6. Why was the hunter lucky?
 - a. He found mud with which he camouflaged himself.
 - b. He found the hippos so easily.
 - c. He found a dying hippo.
 - d. He was not in a hurry to stop searching.

7. Why was it hard to open the mouth of the hippo?
 - a. It was big as a cave.
 - b. The mouth had stuck shut.
 - c. He couldn't get a good grip of the hippos jaws.
 - d. It was full of teeth.

 8. How did the hunter manage to make the mouth remain ajar?
 - a. The teeth settled against each other holding the mouth open.
 - b. He placed a rock between the teeth.
 - c. The mouth remained open due to rigor mortis.
 - d. He had to use all of his strength to keep the mouth ajar.
-
9. How did the hunter manage to make the hippo's mouth remain completely open?
 - a. The mouth was so stiff that it remained open.
 - b. He could hold the mouth open with his back.
 - c. He propped his spear between the jaws.
 - d. He placed his knife between the teeth.

 10. How did the hunter get into the mouth of the hippo?
 - a. He dragged himself along the tongue.
 - b. He walked inside.
 - c. He kicked the obstacles from his way.
 - d. He crawled inside.

Text 3

1. What did Motoko do to her hair as a teenager?
 - a. She cut it waist-long.
 - b. She tried different hairstyles.
 - c. She washed it often to make it shine.
 - d. She brushed it obsessively.

2. What did Motoko do after school days?
 - a. She braided her hair to a single long braid.
 - b. She opened her braid and shook the hair loose.
 - c. She combed her hair to make it look better.
 - d. She chewed on her hair because it made her feel better.

3. What did Motoko study in the yobiko or the extra school?
 - a. Chemistry and English.
 - b. Japanese and English.
 - c. Mathematics and English.
 - d. German and English.

4. At what time did the yobiko end?
 - a. One was free to go after finishing their exercises.
 - b. At nine p.m.
 - c. At ten p.m.
 - d. It depended on the day.

5. What had happened by the time Motoko got off the train?
 - a. The crowd had thinned.
 - b. The train had slowly become packed with people.
 - c. The heavy book bag had started to weigh on Motoko's legs.
 - d. Motoko had changed her school uniform to normal clothes.

6. Where did Motoko put her book bag?
 - a. She forgot the bag into the train.
 - b. Into a locker next to the karaoke bar.
 - c. Being tired she accidentally dropped it to a garbage bin.
 - d. Into a basket attached to her bike.

7. Why didn't Motoko scream when someone grabbed her?
 - a. She knew that no one would hear her in the noise.
 - b. The man covered her mouth with his hand.
 - c. She was too astonished to scream.
 - d. The man told her not to make a sound.

 8. How did Motoko's leg become scratched?
 - a. The man scratched her as they struggled.
 - b. The leg scraped the ground as she fell.
 - c. It hit the pedal of her bicycle.
 - d. Her books fell on her leg.
-
9. What did Motoko realize when she had reached her home door?
 - a. Her waist had started to hurt.
 - b. She had not breathed the whole time.
 - c. She had run at least a mile.
 - d. She had lost her hearing.

 10. What had happened to Motoko's blouse?
 - a. It had become untucked.
 - b. The blouse had been torn.
 - c. It was covered in dirt.
 - d. Nothing happened to the blouse.

Text 4

1. What had the man said about fairies?
 - a. The person sitting next to you could be one of them.
 - b. The fairies like people.
 - c. They often have black dogs as pets.
 - d. The fairies know everything about people.

2. What did the black dog do?
 - a. It chased the people who were coming from the pub.
 - b. It cursed people's souls.
 - c. It blocked the bridge with its backside.
 - d. It talked to people at nights.

3. What did people do to avoid confronting the dog?
 - a. They used a detour or crossed the river at another spot.
 - b. They moved outside only during the summer and stayed inside during the winter.
 - c. They waited for the river to flood because then the dog wouldn't be seen.
 - d. They tried to scare the dog by shouting.

4. How did the priest react when he was asked for help?
 - a. Even the priest was so afraid of the dog that he didn't want to help.
 - b. He preached about going to pubs.
 - c. The priest prayed to God for help.
 - d. He tried to comfort the people.

5. Why were the women sent to the priest?
 - a. Because people knew that the priest had a crush on a beautiful woman.
 - b. They needed to get out of home for a change.
 - c. The women were more devout believers.
 - d. Because women weren't going to the pubs at that time.

6. What did the priest take with him when he was looking for the black dog?
 - a. Nothing.
 - b. A weapon and ammunition.
 - c. One of the village women.
 - d. Holy water and a prayer book.

7. What did the priest do on the fourth day?
 - a. He looked for the black dog day and night.
 - b. He rested for the fourth day was Sunday.
 - c. He held a sermon in which he called the people superstitious.
 - d. He prayed help for the situation for the whole day.

 8. How did the people know the priest had met the dog?
 - a. They went out to investigate what was going on.
 - b. They heard howling that was much louder than that of a normal dog.
 - c. They knew that the dog would show itself at least once every seven days.
 - d. They secretly followed the priest and witnessed the confrontation.
-
9. What was the black dog doing when it was encountered on the next day?
 - a. The dog was lying in the river dead with its legs toward the sky.
 - b. It was chewing on the body of cow it had killed.
 - c. It was chasing people's horses on the pasture.
 - d. It ran away and was never seen again.

 10. What does the storyteller say about people who say that black monster dogs do not exist?
 - a. They are right, because only madmen are superstitious.
 - b. They have never seen such a dog and do not know what they are talking about.
 - c. They should be able to tell fables from true stories.
 - d. They should believe what their elders tell them.

3rd grade questions**Text 1**

1. What did Hera want to do after making love?
 - a. Make love again.
 - b. Go to sleep.
 - c. Have a conversation.
 - d. Go to the face of the earth.

2. What was the question Zeus and Hera ended up with?
 - a. Does a man or a woman fall in love more easily?
 - b. To whom does lovemaking matter the most?
 - c. Which is more genuine, the love of a man or a woman?
 - d. Who takes most pleasure from the act of love?

3. Why would Tiresias know the answer to the question?
 - a. Because he had studied the mating of snakes.
 - b. Because he knew men and women the best.
 - c. Because he had lived both as a man and a woman.
 - d. Because he was the wisest human of all.

4. What did Tiresias do when he encountered the snakes?
 - a. He hit them with a stick he found.
 - b. He started to ponder on the nature of love.
 - c. He walked away so that they wouldn't bite him.
 - d. He grabbed them and threw them to the ground.

5. What were the snakes doing when Tiresias encountered them again?
 - a. They were making love entwined to each other.
 - b. They lied still staring at Tiresias.
 - c. They lunged forward biting Tiresias.
 - d. They slithered away.

6. How was Tiresias like after encountering the snakes for the second time?
 - a. Thoughtful and quiet.
 - b. As manly as he had ever been.
 - c. Old, but wiser than before.
 - d. Disappointed at the twists in his life.

7. How was Tiresias' reaction like when he met the gods?
 - a. Fear made him speechless at first.
 - b. He didn't dare lift his gaze.
 - c. He trembled with fear.
 - d. Terrified he hid behind the bed.

 8. How did Tiresias answer the gods' question?
 - a. He tried to avoid answering directly.
 - b. As plainly as he knew how.
 - c. He tried to lie to please the gods.
 - d. He could not give a direct answer.
-

9. What did Hera do to Tiresias?
 - a. She hit him on the eyes blinding him.
 - b. She punctured his eyes with her fingers.
 - c. She gouged his eyes out with bare hands.
 - d. She forced Tiresias to blind himself.

10. What did Zeus do to Tiresias?
 - a. He cured him with the touch of his finger.
 - b. Zeus could not help him.
 - c. He gave him new eyes.
 - d. He opened up his inner eye.

Text 2

1. What is told about the milieu of the story (time and place)?
 - a. The story is set in a place far away.
 - b. The story is set in a place close by.
 - c. The story is set in a distant time.
 - d. The events of the story took place not so long ago.

2. What did the great hunter think of the medicine man's advice?
 - a. The advice was good, but the hunter wanted more.
 - b. He decided to follow the advice.
 - c. The medicine the witchdoctor offered wasn't enough in the hunter's opinion.
 - d. The medicine man's advice was not good enough for him.

3. What would make the hunter great according to the voice in the dream?
 - a. Killing a hippo.
 - b. Cooking the meat of a hippo.
 - c. Cutting off a hippo's tongue.
 - d. Eating the tongue of a hippo.

4. When did the hunter wake up?
 - a. Early in the following day.
 - b. In the middle of the night, right after the dream.
 - c. Not until days had passed.
 - d. In the evening of the next day.

5. How were the hippos like?
 - a. Docile.
 - b. Scary.
 - c. People-friendly.
 - d. Old.

6. Why was the hunter lucky?
 - a. One hippo had been stuck in the mud.
 - b. He found the hippos so easily.
 - c. He found a dying hippo.
 - d. The hippos had become weaker because of old age.

7. Why was it hard to open the mouth of the hippo?
 - a. It was too big.
 - b. The hunter wasn't strong enough.
 - c. The hunter was too tired.
 - d. The hippo struggled.

 8. How did the hunter manage to make the mouth remain ajar?
 - a. The teeth settled against each other holding the mouth open.
 - b. He placed a rock between the teeth.
 - c. The mouth remained open due to rigor mortis.
 - d. He had to use all of his strength to keep the mouth ajar.
-
9. How did the hunter manage to make the hippo's mouth remain completely open?
 - a. He cranked the mouth open with his spear.
 - b. He pushed it open with as much strength as he could.
 - c. He propped his spear between the jaws.
 - d. He placed his knife between the teeth.

 10. How did the hunter get into the mouth of the hippo?
 - a. He dragged himself along the tongue.
 - b. He walked inside.
 - c. He kicked the obstacles from his way.
 - d. He crawled inside.

Text 3

1. What did Motoko do to her hair as a teenager?
 - a. She cut it waist-long.
 - b. She tried different hairstyles.
 - c. She washed it often to make it shine.
 - d. She brushed it obsessively.

2. What did Motoko do after school days?
 - e. She braided her hair to a single long braid.
 - f. She opened her braid.
 - g. She polished her hair.
 - h. She shook her braid.

3. What did Motoko study in the yobiko or the extra school?
 - a. Chemistry and English.
 - b. Japanese and English.
 - c. Mathematics and English.
 - d. German and English.

4. When did Motoko get home on the days she went to yobiko?
 - a. It varied depending on the ending time of the yobiko.
 - b. Nine p.m.
 - c. Ten p.m.
 - d. It varied depending on when she finished her homework.

5. What had happened by the time Motoko got off the train?
 - a. The crowd had thinned.
 - b. The train had slowly become packed with people.
 - c. The heavy book bag had started to weigh on Motoko's legs.
 - d. Motoko had changed her school uniform to normal clothes.

6. Where did Motoko drop her book bag?
 - a. To the bicycle parking lot.
 - b. To a locker made of wire.
 - c. To a red garbage bin.
 - d. To the basket of her bicycle.

7. Why didn't Motoko scream when someone grabbed her?
 - a. Because the situation seemed so unreal.
 - b. She tried but she couldn't make any sound.
 - c. She was too astonished.
 - d. She was too busy struggling.

 8. How did Motoko's leg become scratched?
 - a. The man scratched him as they struggled.
 - b. The leg scraped the ground as she fell.
 - c. It hit the pedal of her bicycle.
 - d. Her books fell on her leg
-

9. Why did Motoko have to bend over at her home door?
 - a. Her waist had started to hurt.
 - b. She had not breathed during the escape.
 - c. She had run a mile.
 - d. Her head had started to ache.

10. What had happened to Motoko's blouse?
 - a. It had become untucked.
 - b. The blouse had been torn.
 - c. It was covered in dirt.
 - d. Nothing happened to the blouse.

Text 4

1. How were the fairies like according to the man?
 - a. They often sit around people.
 - b. The fairies like people.
 - c. You could think they're people.
 - d. The fairies know everything about people.

2. What did the black dog do?
 - a. It chased the people who were coming from the pub.
 - b. It cursed people's souls.
 - c. It blocked the bridge with its backside.
 - d. It talked to people at nights.

3. What did people do to avoid confronting the dog?
 - a. They moved outside only during the summer.
 - b. They waited for the river to flood.
 - c. They changed their route.
 - d. They moved outside only during the winter.

4. How did the priest react when he was asked for help?
 - a. Even the priest was so afraid of the dog that he didn't want to help.
 - b. He invited people to a sermon.
 - c. He told people to stay in the pubs.
 - d. He preached about going to pubs.

5. Why were the women sent to the priest?
 - a. Because people knew that the priest had a crush on a beautiful woman.
 - b. They needed to get out of home for a change.
 - c. The women were more devout believers.
 - d. Because women weren't going to the pubs at that time.

6. How did the priest arm himself when looking for the black dog?
 - a. He didn't take anything with him.
 - b. The storyteller didn't know but he assumed that the priest was armed.
 - c. He took holy items with him.
 - d. He had holy water and a prayer book.

7. What did the priest do on the fourth day?
 - a. He looked for the black dog day and night.
 - b. He rested for the fourth day was Sunday.
 - c. He held a sermon in which he called the people superstitious.
 - d. He prayed help for the situation for the whole day.

 8. How did the people know the priest had met the dog?
 - a. They went out to investigate what was going on.
 - b. They heard howling that was much louder than that of a normal dog.
 - c. They knew that the dog would show itself at least once every seven days.
 - d. They secretly followed the priest and witnessed the confrontation.
-
9. How was the black dog met on the next day?
 - a. The dog was found dead in the river.
 - b. The dog was levitating in air above the river.
 - c. The dog was diving in the river.
 - d. The dog had moved to the river bank close to the sea.

 10. What does the storyteller say about people who say that black monster dogs do not exist?
 - a. They are right, because only madmen are superstitious.
 - b. They have never seen such a dog and do not know what they are talking about.
 - c. They should be able to tell fables from true stories.
 - d. They should believe what their elders tell them.

Appendix 4. The SE questionnaire in Finnish

Ympyröi mieleisesi vaihtoehto.

Kertojan puhetyyli sai hänet kuulostamaan:

1. vähä-älyiseltä	1	2	3	4	5	6	7	älykkäältä
2. sivistymättömältä	1	2	3	4	5	6	7	sivistyneeltä
3. matalasti koulutetulta	1	2	3	4	5	6	7	korkeasti koulutetulta
4. epäystävälliseltä	1	2	3	4	5	6	7	ystävälliseltä
5. huumorittomalta	1	2	3	4	5	6	7	humoristiselta
6. epärehelliseltä	1	2	3	4	5	6	7	rehelliseltä

Kertoja puhui:

7. liian hitaasti	1	2	3	4	5	6	7	liian nopeasti
-------------------	---	---	---	---	---	---	---	----------------

Kertojaa oli:

8. helppo ymmärtää	1	2	3	4	5	6	7	vaikea ymmärtää
--------------------	---	---	---	---	---	---	---	-----------------

Videon näkeminen oli kuullun ymmärtämisen kannalta:

9. hyödytöntä	1	2	3	4	5	6	7	hyödyllistä
---------------	---	---	---	---	---	---	---	-------------

Appendix 5. The SE questionnaire in English

Circle the alternative that corresponds to your opinion.

The storyteller's style of speech made them sound:

- | | | | | | | | | |
|------------------------|---|---|---|---|---|---|---|-----------------|
| 1. of low intelligence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | intelligent |
| 2. unsophisticated | 1 | 2 | 3 | 4 | 5 | 6 | 7 | sophisticated |
| 3. of low education | 1 | 2 | 3 | 4 | 5 | 6 | 7 | highly educated |
| 4. unfriendly | 1 | 2 | 3 | 4 | 5 | 6 | 7 | friendly |
| 5. humorless | 1 | 2 | 3 | 4 | 5 | 6 | 7 | humoristic |
| 6. dishonest | 1 | 2 | 3 | 4 | 5 | 6 | 7 | honest |

The storyteller spoke:

- | | | | | | | | | |
|-------------|---|---|---|---|---|---|---|----------|
| 7. too slow | 1 | 2 | 3 | 4 | 5 | 6 | 7 | too fast |
|-------------|---|---|---|---|---|---|---|----------|

The storyteller was:

- | | | | | | | | | |
|-------------------------------------|---|---|---|---|---|---|---|--------------|
| 8. easy to understand
understand | 1 | 2 | 3 | 4 | 5 | 6 | 7 | difficult to |
|-------------------------------------|---|---|---|---|---|---|---|--------------|

For listening comprehension seeing the video was:

- | | | | | | | | | |
|------------|---|---|---|---|---|---|---|--------|
| 9. useless | 1 | 2 | 3 | 4 | 5 | 6 | 7 | useful |
|------------|---|---|---|---|---|---|---|--------|