

Hilda Perander

**”ONNELLISINA HE LIITÄVÄT LIAANILLA
HÄNNÄT VIUHUEEN”**

Shifting between frames in audio described
framed narrative

ABSTRACT

Hilda Perander : “Onnellisina he liitävät liianilla hännät viuhuen”: Shifting between frames in audio described framed narrative

Bachelor's thesis

Tampere University

Bachelor's Programme in English Language, Literature and Translation

April 2020

The purpose of this study is to investigate how audio description supports narration in a story that is built around a framed narrative structure. A framed narrative is a form of storytelling where the characters of one story tell each other a story. In this study, I use a theory on framed narrative as presented by Katharine Young to divide the narrative shifts into smaller pieces and to determine if a shift is interrupted or not.

Audio description is an accessibility practice where visible material is described in an audible form for blind or visually impaired people. In this study, I use the categories presented by Anna Marzá Ibañez to classify the informational content in the description. The classification states that audio description answers to the following questions about the material: where, when, what, who, and how. A third central theory is Christopher Taylor's theory on multimodal transcription which I adapted to depict audio described material for this thesis.

I studied an episode from a Finnish-British children's TV show *Moominvalley* called Muumipapan kalajuttu, which contains 12 shifts within the narrative. I inspected the pieces of audio description that occur in the instance of a shift when the narrative moves to another dimension physically or temporally. I classified these pieces of audio description based on the amount of information they contain and the interruption status and the direction of the shift to which they relate.

The research shows that the direction of a shift somewhat affects the amount of information in a piece of audio description. When the narrative moves from an inner story to an outer story, i.e. from a story to a conversation between the characters, a piece of audio description relating to that shift is often shorter and contains less types of information than in situations where the narrative moves from conversation to a story. The interruption status of a shift seemed to have little effect on audio description based on this study.

The results of this study indicate phenomena in a singular occasion and cannot be generalized to all of audio described material. Future research could include studies on Finnish audio described children's media in general. Audio description for Finnish movies and TV-shows has been done for years by a small number of professionals, and research on the field is sparse. It could be beneficial to examine the description tendencies of different audio describers and compare them to each other and to existing guidelines to get a better picture on the situation and the level of Finnish audio description.

Keywords: audio description, framed narrative, children's media, accessibility

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TIIVISTELMÄ

Hilda Perander : “Onnellisina he liitävät liianilla hännät viuhuen”: Shifting between frames in audio described framed narrative

Kandidaatintutkielma

Tampereen yliopisto

Englannin kielen, kirjallisuuden ja kääntämisen kandidaattiohjelma

Huhtikuu 2020

Tämän tutkimuksen tarkoituksena on selvittää, miten kuvailutulkkaus tukee kerrontaa tarinassa, joka rakentuu kehyskertomuksen ympärille. Kehyskertomus on kerronnan muoto, jossa jonkin tarinan hahmot kertovat toisilleen tarinaa. Käytän tutkimuksessani Katharine Youngin esittelemää kehysteoriaa, jonka mukaan jaan kerronnan väliset siirtymät pienempiin osiin ja määrittelen, keskeytyvätkö ne.

Kuvailutulkkaus on saavutettavuuden muoto, jossa näkyvää materiaalia kuvaillaan sanalliseen muotoon sokeille tai näkövammaisille ihmisille. Käytän tässä tutkimuksessa tulkkeiden sisältävän informaation luokitteluun Anna Marzà Ibañezin kategorioita, joiden mukaan kuvailutulkkaus vastaa kuvailevasta materiaalista seuraaviin kysymyksiin: missä, milloin, mitä, kuka ja miten. Kolmas keskeinen teoria on Christopher Taylorin teoria multimodaalisesta transkriptiosta, jota sovelsin kuvailutulkatun materiaalin kuvantamiseen tätä tutkimusta varten.

Tutkin suomalaisbritannialaisen lasten TV-sarjan *Muumilaakso* jaksoa Muumipapan kalajuttu, jossa on 12 siirtymää kertomusten välillä. Tarkastelin kuvailutulkkeita, jotka sijoittuvat siirtymissä siihen kohtaan, kun kerronta siirtyy toiseen ulottuvuuteen joko fyysisesti tai ajallisesti. Luokittelin nämä kuvailutulkkeet niiden informaation määrän sekä niihin liittyvän siirtymän keskeytyneisyyden ja suunnan mukaan.

Tutkimuksesta selviää, että siirtymän suunnalla on jonkin verran vaikutusta kuvailutulkkeen sisältämän informaation määrään. Kun kerronta siirtyy sisätarinasta ulompaan tarinaan, eli tarinasta hahmojen väliseen keskusteluun, siirtymään liittyvä tulke on usein lyhyempi ja sisältää vähemmän erityyppistä informaatiota kuin tilanteissa, joissa kerronta siirtyy keskustelusta tarinaan. Siirtymän keskeytyneisyydellä ei tämän tutkimuksen perusteella näytä olevan vaikutusta kuvailutulkeeseen.

Tämän tutkimuksen tulokset osoittavat ilmiöitä yksittäisestä tapauksesta, eikä niitä voi yleistää kaikkeen kuvailutulkattuun materiaaliin. Jatkotutkimuksia voisi tehdä yleisesti suomalaisesta kuvailutulkatusta lasten mediasta. Suomessa elokuvien ja tv-ohjelmien kuvailutulkkauksia on vuosia tehnyt vain pieni joukko ammattilaisia, ja alan tutkimusta on niukasti. Voisikin olla hyödyllistä tarkastella eri kuvailutulkkien tulkaustapoja sekä vertailla niitä keskenään ja olemassa oleviin ohjeistuksiin, jotta saataisiin parempi kuva suomalaisen kuvailutulkkauksen tilanteesta ja tasosta.

Asiasanat: kuvailutulkkaus, kehyskertomus, lasten media, saavutettavuus

Tämän tutkielman alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

Table of contents

- 1 Introduction 1
- 2 Theoretical framework 3
 - 2.1 Audio description 3
 - 2.2 Frame theory 4
 - 2.3 Multimodal analysis 5
- 3 Material and method 7
- 4 Analysis 11
 - 4.1 Shifts with less information 12
 - 4.2 Shifts with more information 14
 - 4.3 Comparison of the two groups 16
 - 4.4 Division according to the interruption status 18
- 5 Conclusion and discussion 19
- 6 References 22
 - 6.1 Research data 22
 - 6.2 Other references 22
- 7 Appendices 24

1 Introduction

Audio description (AD) is a practice where visual information is transferred into auditive information. The main audience for AD is blind and visually impaired people. Target materials of AD usually include film, TV, art, and theatre. This thesis focuses on the AD of an episode of *Muumilaakso*, a Finnish-British 3D-animated children's TV production about Moomins. In the Finnish context, Moomins are a beloved cultural aspect and have been popular among children and their parents for decades. With the inclusion of AD in the newest production, blind and visually impaired Finnish children can now enjoy the story in a way they might not have been able to enjoy previously.

The theories this thesis uses for its analysis are Katharine Young's *frame theory* and the guidelines for audio description by Anna Marzà Ibañez. The material is collected by adapting Christopher Taylor's theory on *multimodal transcription*.

Young's (2010) theory introduces a structure proposal for a framed narrative. A framed narrative is a way of storytelling where characters in a story tell each other a story. Young's theory describes four frames that precede a story and four frames that follow it, and each of them are surrounded by the conversation from which the story emerges. In this thesis, I will apply the theory to the material and inspect more closely the situations where some frames are missing due to an interruption in the structure.

Marzà Ibañez (2010) details a guideline for AD students on what types of information AD for film should contain, when to describe in relation to other sounds, and how to present the description. This thesis focuses on the guidelines for what to describe and compares them to the findings from the material.

Taylor's (2016) theory on multimodal transcription details a way to transcribe multimodal information into a format where it can be more easily studied. A multimodal transcription consists of rows and columns, where each row represents a moment of time in the multimodal material and each column describes the material in its visual, kinesic, and auditive aspects. In Taylor's theory, one column is reserved for subtitles, as his approach is to inspect how the transcript method might aid translators in creating more accurate subtitles, but in the interests of this thesis, that column is substituted for a column detailing the AD in each scene.

The material chosen for this research is picked from the eighth episode of the first season of *Muumilaakso* called Muumipapan kalajuttu (2019), and the AD for it was made by Anu Aaltonen, Hannele Antikainen and Carita Lehtniemi. This episode was chosen specifically because it includes a framed narrative which structures the episode into a drastically different direction compared to the other episodes in the first season. The episode contains multiple narrative shifts to and from stories that the characters tell each other. Many of these shifts include a striking visual element, a freeze image, which is a clear indicator to a viewer that something unusual is about to transpire. An audience member who only relies on hearing or partial sight, however, will not receive the same clue.

This possible unbalance between visual and auditive information made me interested in studying the occurrence further and led me to the formulation of the research questions for this thesis. How does the AD support the narrative shift from the story to the conversation and back? What differences and similarities can be drawn from the material based on the factors present in the shifts, such as the direction of the shift or its interruption status? I hope to find answers to both of these questions during the research.

There is little previous research on the topic of audio description. In Finland, AD has been the topic of one doctoral dissertation (Hirvonen 2014) and six master's theses (Heinemaa 2017, Nuutinen 2014, Pesonen 2015, Reiman 2017, Vaittinen 2018, Valli 2017), which were all published during the last seven years. Three of them focus on AD in film (see Hirvonen 2014, Reiman 2017, Vaittinen 2018), but none of them focus on narrative as a central topic. Similarly, there is little research on AD for children specifically (see Palomo López 2010). According to Louise Fryer, audio describing material for children means using different, often opposing, strategies compared to AD for a general audience, such as repetition and playfulness (Fryer 2016, 114–15).

This thesis contains seven chapters including this introductory chapter. The following chapter describes what AD is and focuses more on the other two theories. The third chapter contains a description of the episode as research material and the method to collect the pieces of AD to study further. The fourth chapter describes the analysis of the material collected in the previous chapter. The fifth chapter concludes the research, attempts to arrive at an answer to the research questions, and critically analyses the part of AD in the research material. The last two chapters contain a bibliography and references as well as a list of appendices.

2 Theoretical framework

2.1 Audio description

Audio description attempts to convey as much essential visual information in an audible form as is possible and required. It is characterized by time restrictions that arise from the material that is being described. Ideally, AD should not overlap with dialogue, meaningful music, or other sounds already present in the material. Therefore, in a situation with very little available audible space, a describer must decide what the most important factor in a scene to describe is, oftentimes at the expense of other details.

AD is aimed at being neutral and lacking any attitude from the describer. Different materials require different describing styles, and it is therefore difficult, if not impossible, to present strict guidelines that would cover all of AD. Most guidelines cover the very surface layer of AD and do not give specific, detailed instructions (see Kuvailutulkkaustoimikunnan yleinen kuvailutulkkausohje, 2012). Some of the guidelines are also contradictory to each other, such as in the case of whole sentences. The Finnish organization for audio description, Kuvailutulkkaustoimikunta (2012), recommends using whole sentences whenever possible, whereas Fryer (2016, 55) compares the regularity of incomplete sentences in AD to common everyday speech. The opposing viewpoints may be the result of one set of guidelines trying to appeal to all of AD and another to just one style of AD, such as describing only sculptures or only TV-shows. See Heinemaa (2017) for more discussion on AD guidelines.

Among other guidelines regarding the entire process of producing an AD script, suggestions have been made as to what type of information AD does or should contain. Philip Piety (2004) investigated a corpus of audio described movies and created a list of the common information types that appeared in the AD. In the same study, Piety introduces the term *utterance* (ibid. 12) among others to describe a unit of spoken language, which I will use in this study to refer to the pieces of AD.

In another guideline, Marzà Ibañez (2010) lists a suggestive order for information when describing a film, of which the relevant categories to this study are: *where*, *when*, *what*, *who*, and *how* (ibid. 147). The *where* relates to a location of characters in relation to each other or their surroundings. The *when* relates to the time in the film. The *what* relates to the action in a scene. The *who* relates to the characters: their physical appearance, expressions, and clothing.

For this thesis, I will expand the *who*-category to include non-humanoid creatures and other objects that appear as an agent in a scene. Finally, the *how*-category describes a manner of action. In this thesis, I will use the categorization presented by Marzà Ibañez to determine the amount of information in an utterance. The more categories present in an utterance, the more information it has.

2.2 Frame theory

Katharine Young’s frame theory (2010) details the frames which surround a story and frame it within a conversation. A framed narrative is a story within a story: a story that characters in a story tell each other during ordinary conversation. According to Young’s theory, each instance of a framed narrative contains eight types of frames that set the story apart from the conversation from which the story emerges. Four frames precede the story, and four frames follow it. Each frame belongs to one of three conversational realms. The *realm of conversation* contains the conversational aspect of the surrounding story, where the characters converse with each other. The *storyrealm* begins when one character begins to tell a story while the discourse is still in the realm of conversation and ends as the discourse moves from discussing the story to ordinary conversation. The *taleworld* contains the story itself, and it might be completely removed from the characters and the location of the surrounding story. A more detailed view on the realms and the frames is in Table 1.

Table 1. Parts of frame theory

Realm	Frame	Explanation
Realm of conversation	(Conversation)	Is the discourse between the characters in the surrounding story.
Realm of conversation	Preface	Informs the listeners that a story is about to take place.
Storyrealm	Opening	Shifts the discourse from conversation to narrative.
Storyrealm	Orientation	Provides essential information without which it would be difficult to follow the story.
Taleworld	Beginning	Contains parts of the story where it begins.
Taleworld	(Story)	Is the story.
Taleworld	End	Contains parts of the story where it ends.

Storyrealm	Closing	Shifts the discourse from narrative to conversation.
Storyrealm	Evaluation	Contains the listeners' reactions to the story.
Realm of conversation	Coda	Informs the listeners that the story is over and that they can now take turns in the discourse.
Realm of conversation	(Conversation)	Is the discourse between the characters in the surrounding story.

Even though *conversation* and *story* are not frames per se but the narrative pieces that are being framed by the other frames, they are included in the table in parentheses for clarity's sake.





For this thesis, I define a *shift* as a section of discourse that contains the frames between conversation and story in either direction. This thesis focuses on the shifts that occur from the taleworld to the realm of conversation, especially in places where frames are skipped due to an interruption of the story. This study also considers the visual content in the story-telling aspect. Thus, something that is not conveyed by audio but by visuals can be called a frame. A surprised look on a listener's face can be called evaluation since it conveys a meaningful reaction to a story.

In this thesis, I mention a *previous similar situation* in relation to some stories and pieces of conversation. With the term, I refer to a previous story or conversation that occurred in the same location with the same characters as the situation in question. This is done to determine if some information is already known to the audience from an earlier story or conversation and therefore would not require a description when the information is shown again.

2.3 Multimodal analysis

Christopher Taylor's adaptation on *multimodal theory* (2016) lines out a method of transcribing audio-visual material with a specific interest in translation. The basic transcript is a table with four columns. Each row in the table represents a moment in the audio-visual material at predetermined intervals, which in this study is two seconds. Table 2 below is an excerpt of the multimodal transcription where the method has been adapted for this study. The transcript is of shift 3 which occurs between the second and the third row. For this thesis, the theory has been adapted from Taylor's method. A column has been added to the beginning to detail the time, and subtitles have been changed to audio description in the last column.

Table 2. Example of multimodal transcription

Time, row	Visual frame	Visual image and kinesic action	Soundtrack	Audio description
06:36 1		Pappa and Moomintroll are lying on the beach. Moomintroll sits on a pool of green slime with a large wooden box behind him. Pappa turns to look at Moomintroll with annoyance and talks.	Pappa: “Päästit sitten vavasta irti.”	-
06:38 2		Pappa looks at Moomintroll, who looks back and blinks. Neither of them moves.	-	“Kotipihalla.”
06:40 3		Cut to Pappa in the yard. He sits in front of the box and talks with a neutral expression.	Pappa: “Ja niin tarunhohtoinen mamelukki”	-
06:42 4		Pappa talks. Cut to Snorkmaiden, Mamma, Moomintroll and My sitting on the ground with surprised expressions.	Pappa: “pääsi karkuun.”	-

The first column details the time stamp of the row as well as an identifying number for each row. The second column contains an image of the scene, here depicted as traced drawings of the screenshots. The third column contains a written description of the image and the movement that occurs during the interval. The fourth column contains the soundtrack and the spoken lines during the interval. In Taylor’s method, the fifth column would contain the subtitles. Since the research material contains no subtitles but does contain an audio description track, the last column has been changed from subtitles to audio description. Although it could be argued that audio description is merely a part of the soundtrack of the episode such as the music or the spoken lines and thus would belong in the soundtrack column, dividing them into their own columns is better suited for the interests of this thesis as audio description is a separate mode of communicating information.

3 Material and method

In this chapter, I will introduce the material chosen for this thesis. Then, I will detail the process of extracting the pieces of AD to be studied further.

In the episode, Moominvalley experiences a drought that has destroyed all the crops. Moomintroll and Moominpappa are on a fishing trip to get food for the rest of the household: Moominmamma, Snorkmaiden, and Little My. The fishers return with a box full of seeds, and Moominpappa sets out to tell a story to the others about their fishing trip. The action-packed story is revealed to be false and exaggerated, and Moomintroll tries to recount a more faithful depiction of their fishing trip. Moominpappa replies that his son might not be as captivating a storyteller as he himself is. During the night, the seeds spill out of the box and, once the morning arrives, start to grow in an unnatural speed. Soon the plants reach the roof of the Moominhouse and have begun growing indoors as well, turning the house into a jungle. Moomintroll and Little My venture around the house, saving Snorkmaiden on the way. Abruptly, the narrative cuts to all of the characters sitting in a dusty attic, seemingly detached from the previous action. It turns out that Moomintroll is merely telling the others a story of how he would save the others from the perils of monster plants, perhaps to prove to his father that he is a good storyteller after all. For the remainder of the episode, the narrative cuts between the story and the reality multiple times. In the last scene, the audio from the reality-part overlaps the visuals of the story-part, which creates an interesting narrative conclusion to the story and the episode overall. A more detailed timeline of the episode is included in Appendix 1.

I used multimodal transcription as introduced by Christopher Taylor (2016) to transcribe the scenes that included shifting between the narrative frames. I transcribed eight scenes which contained a total of 156 seconds of material, 78 rows of transcript and 12 shifts¹ between story and conversation. For the purposes of this thesis, I chose an interval of two seconds for each

¹ Arguably, there are two more shifts in the material, but I have chosen to exclude them from this study. Shift “0.5” occurs between the very beginning of the episode and the first part of the fishing-story, but it is realised in the form of the opening sequence to the episode. The opening sequence exists on a level above the narrative and the content of the episode and is therefore not comparable to the other shifts that happen within the narrative and by the characters’ initiative. Shift “5.5” is excluded for reasons of clarity. It is clear that shift 5 ends in the realm of conversation and that the narrative preceding shift 6 is within the taleworld, but during the time between the two shifts, there is no clear indicator of a conversation-to-story shift. As this is left unclear for the viewers, it is a reasonable choice in the describer’s part to leave it undescribed in the AD as well.

transcription. This was a sufficiently small interval to capture the shifts which often included a still image or a blackout over multiple seconds. A smaller interval, such as one second, would have been unnecessarily small since I was less interested in the specifics of the order of the spoken lines and the AD. In the next paragraph, I will explain how I transcribed the scenes.

First, I familiarized myself with the episode well enough to be able to identify the shifts between the narrative frames. Then, I marked all the timestamps of such scenes in a table. For each scene, I watched the scene through a few times and determined where to begin and end the transcript. I began the transcript just before the shift where either a spoken line or the AD establishes the current situation or place. This could be either in the tale-world or in the realm of conversation. To illustrate, the transcript in Table 2 begins with Moominpappa's final line within the story and the taleworld. I ended the transcript to the first bit of information that establishes the setting into the opposing realm after the shift. In Table 2, that information is the AD informing the audience of a physical shift as well as the visual shift into the home yard.

After determining the scenes to transcribe, I specified the time segments into a five-column table. The first column contains the beginning time of each segment as minutes and seconds of the episode, and underneath it is a running number of the frames per each scene. Then, I watched the scene piece by piece, wrote a short description of each two-second segment, and screenshotted the images. I added the images into the second column of the transcript, and for Table 2, I traced the relevant screenshots on an image processing program. After this, I filled in the AD column into the fifth column. If any piece of the AD continued onto the following segment, it is split apart into the correct rows. If the segment had no AD, the row is marked with a dash. Then, I added the spoken lines and sound effects into the fourth column. Lastly, I wrote a description of the action in each segment into the third column.

I analysed each row and indicated which part of the narrative is happening in each row based on the description of each frame in Young's material. Each frame is indicated by a spoken line by the characters, something happening in the visual space, a piece or audio description, or a combination of these. I wrote down what the piece of information was that was indicating a frame or story, with close attention to the parts where AD conveyed the frame. Each shift is also connected to a near-occurring piece of AD that describes a new location, an action that

happens in the new location, or both. The material contains twelve shifts that I will be studying further.

After this analysis, I extracted only the pieces of AD, or the utterances, that relate to a shift, which I have listed in the Table 3 below. Each utterance was then categorized into one of two classes according to the following three criteria: its interruption status, its direction, and its amount of information. The Finnish utterances in Table 3 are as they appear in the material, and below each one is my own translation of the utterance into English.

Table 3. The utterances with their translations, interruption status, direction, and amount of information

#	Utterance (English translation) [all my translations]	Interruption status	Direction	Information, # of types
1	Istuvat kotipihalla. ([They] sit at the home yard.)	interrupted	story-to-conversation	less, 2
2	Mamelukki alkaa vetää venettä kiihtyvällä vauhdilla. (Mamelukki begins to pull the boat at an increasing speed.)	interrupted	conversation-to-story	more, 3
3	Kotipihalla. (At the home yard.)	noninterrupted	story-to-conversation	less, 1
4	Ovat merellä onkimassa. ([They] are fishing at the sea.)	noninterrupted	conversation-to-story	less, 2
5	Pappa kuorsaa vatsallaan laatikon edessä. (Pappa snores on his stomach in front of the box.)	interrupted	story-to-conversation	more, 4
6	Talon ullakolla. (In the attic of the house.)	interrupted	story-to-conversation	less, 1
7	Onnellisina he liitävät liaanilla hännät viuhuen. (Happily they glide on a vine, tails swishing.)	interrupted	conversation-to-story	more, 3
8	Ullakolla. (In the attic.)	interrupted	story-to-conversation	less, 1

9	Pappa, Mamma ja Niiskuneiti perääntyvät portaissa puskapedon hyökätessä heitä kohti. (Pappa, Mamma, and Snorkmaiden back down in a staircase as a plant beast attacks towards them.)	interrupted	conversation-to-story	more, 3
10	Ullakolla. (In the attic.)	noninterrupted	story-to-conversation	less, 1
11	Köynnökset tunkeutuvat ikkunoiden läpi. (Vines push through the windows.)	noninterrupted	conversation-to-story	more, 3
12	Ilta pimenee Muumilaaksossa ja kuu nousee taivaalle. (The evening darkens in Moominvalley and the Moon rises on the sky.)	noninterrupted	story-to-conversation	more, 4

A shift's interruption status was defined as follows: I compared each scene to Young's theory to see if any frames were missing from the order. If any frames were missing, the shift was denoted *interrupted*, meaning that something happens in the narrative that interrupts the regular flow of storytelling. If no frames were missing, the shift was denoted *noninterrupted*. In total, there are seven interrupted shifts and five noninterrupted shifts. The most often skipped frames were *coda*, *end*, *opening*, and *beginning*.

The direction of a shift is either from conversation towards a story or from a story towards conversation. While the material chosen for further research has no parts that could be labelled conversation, as there is no discourse between characters that does not pertain directly to the story-telling aspect of the situations, I marked the directions as *conversation-to-story* when they contained any frames from preface to beginning, and *story-to-conversation* when they contained any frames from end to coda. In total, five shifts are conversation-to-story and seven shifts are story-to-conversation.

The amount of information relates to the number of information types present in an utterance. I noted that this varied greatly; some utterances contained only a one-word-description of a new location, while others contained a detailed description of an action, an agent, a manner of doing, and a new location. A shift has *less information* if the description contained one or two types of information. A shift has *more information* if the description contained three or four types of information. The utterances in this study contained all categories at least once, and each of them contained at least one category. There were no utterances that had all five types

of information present. In total, six shifts had less information and six shifts had more information.

The numbers on their own seem rather even but studying how the categories divide further gives us more insight. Table 4 divides the shifts according to the amount of information in each utterance.

Table 4. Division of the shifts by the amount of information

Less information			More information		
#	Interruption	Direction	#	Interruption	Direction
1	int	st-con	2	int	con-st
3	non	st-con	5	int	st-con
4	non	con-st	7	int	con-st
6	int	st-con	9	int	con-st
8	int	st-con	11	non	con-st
10	non	st-con	12	non	st-con

A detailed report on all of the information types present in the utterances is included in Appendix 2.

4 Analysis

As we can see from Table 4, shifts have less information mostly when the narrative shifts from a story to conversation. Conversely, shifts have more information mostly when the narrative shifts from conversation to a story. However, shifts 4, 5, and 12 oppose these trends in their respective groups. The interruption status seems to have less effect on the amount information of a piece of AD, with similar numbers in both groups and no clear division when compared to the direction of the shifts. In my analysis, I attempt to uncover what is common with the majority of the shifts in both groups by examining the situations of the shifts and investigating the commonalities and the differences in the narrative in each case. First, I will analyse shifts with less information and the shifts with more information as their own groups, and then, I will compare the two groups. Lastly, I will inspect the differences in the AD when divided according to the shifts’ interruption status.

4.1 Shifts with less information

The six shifts with less information contain only two information types, *where* and *what*. Five of them shift from a story to conversation. During the shifts, the changing aspects include the location, the characters, the time, and the action. In the shifts with less information, though, some of these aspects are already known to the audience.

Location, or the *where*, is described in all of the shifts, with four shifts having information only about the location, and it is known in five of the shifts. Shifts 1 and 3 move the action to the yard in front of the Moomin house, from where the episode started and which has been shown to the viewers. The AD does not explicitly name the location as a yard before shift 1, but it mentions Moominmamma and Snorkmaiden worrying about a vegetable plot which can be assumed to be near their home. Shift 4 is the only shift with less information that moves the narrative from conversation to a story, which occurs at the sea. The location is the same as in the previous similar situation, so a one-word description suffices. Shift 6 introduces a new location, the attic of the house, which is not yet familiar to the audience. Therefore, it is described in relation to the building wherein the majority of the episode has been set. Later, in shifts 8 and 10, the attic is described just by mentioning the room, since the location is already known to the audience.

The episode has a limited cast of characters who are present in different scenes. In shifts 1 and 3, the characters present in the new situation are known to the audience as they have all been named by the describer or the other characters by the time of the first shift and are known to be in the location mentioned in the AD. The characters in the story after shift 4 have been established earlier in the episode in the story Moominpappa tells the others. As all or most characters are either known or the same as in a previous similar scene, it is not required to mention them all by name. The Finnish AD takes this one step further by not even saying the pronoun “they” which would be grammatically required. In AD, however, the use of incomplete sentences is a common tactic to save time (Fryer 2016, 55). These two reasons, familiar characters and the common use of incomplete sentences, have, most likely, lead to the *who* being left unanswered in these shorter pieces of AD.

However, in the rest of the shifts, the characters in the scenes are not known. In shift 6, it is not clear who all are present in the attic. Only Moomintroll, Little My and Moominpappa

speak in the brief conversation between shifts 6 and 7, and the AD does not mention that two other characters are present in the attic even on their next appearance in shift 8. They do not speak until shift 10, when the non-seeing audience gets to acknowledge them for the first time. Leaving the *who* out of the AD might have been due to time preservation or a general assumption that when specific characters are not mentioned, the action pertains to the entire cast of characters in the episode. The latter, however, is a faulty tactic and can lead to unnecessary ambiguity. Another reason for not mentioning which characters are present in a scene might be because they remain silent during the scenes and do not act in any noticeable way. Therefore, knowledge of their presence might be deemed non-essential for the audience to understand what is happening in the scene as these characters do not directly contribute to the action.

A distinctive feature in the contents of the utterances is the *what*-category. Only two of the six shifts with less information contain a description of action, whereas all six shifts of more information contain the *what*-category. Why is there so little need to know what is happening in the shifts that mostly occur from a story to conversation? The answer might lie in the inverse of the reason why audio description is needed in the first place: to describe visual information, or to not describe audible information.

Half of the shifts from a story to conversation – shifts 1, 6, and 8 – happen because a character interrupts the storyteller; an evaluation occurs earlier in the order of frames than it is supposed to occur. Without the interruption, the story would continue until the storyteller is finished and a natural shift back to the conversation would happen later. The manner of interruption in all these three cases is a spoken objection to the content of the story. Speaking, by definition, is an audible action, and so is any spoken conversation. Therefore, when a new action in a new scene is specifically conversation between characters, it is inherently audible and does not require a description to a hearing audience. The *what* remains unanswered in the AD; this stands true for shifts 6, and 8 as well as the non-interrupted shifts 3 and 10, as they contain either closing or evaluation that is spoken. Shift 1 contains a verb, *sitting*, which can hardly be called a description of action and instead understood as locating the characters in their surroundings, but the audible act of conversation is still left undescribed.

Overall, the situations in which these shifts occur and which the utterances describe mostly contain something that is already known to the audience or an action that requires no

description, which results in less information to describe. Next, I will analyse the shifts with more information.

4.2 Shifts with more information

The six shifts with more information contain all five types of information in different configurations. The more common direction of a shift is conversation to a story with four utterances with more information in them having this direction. A common thread with the shifts with more information seems to be that in all of them, something novel happens after the shift.

All six shifts contain information about *what*, an action, and in all of them, the action is something new. In none of them is the action something that has been shown before in the episode, and therefore, they all require a description.

Audibility is a central factor in all of the descriptions of action. In shifts 2, 5, and 11, the action is joined by audio that might be either unidentifiable or difficult to connect to a specific source. Shift 2 contains a sound of splashing water, which in itself does not make it clear that a giant fish is pulling a boat. Shift 5 contains audio from Moominpappa snoring before Moomintroll gets to finish his story, but the audio could be mistaken for the other sounds of struggle from pulling the large box moments before. Shift 11 has arguably the clearest sound of these, a shattering glass, but the reason for it is unidentifiable from audio only. These sounds that are tied to a visible action but not immediately recognisable require a description as it is not clear from the audio what causes the sound. Shifts 7, 9, and 12, however, contain non-audible action, such as backing down in a staircase or the Moon rising on the sky. As they are purely visible, they require a description as well. The ambiguous audibility or non-audible action altogether lead to the action requiring a description in each of these shifts, contributing to overall more information in them.

All five instances of *who* in this material occur in the shifts with more information. In shifts 2, 5, 9, and 11, the agent in the situation is someone or something specific that has to be clarified to the audience. Often in these scenes, other characters are present, but only one or some characters' action is described, requiring these characters to be named. This is the case in shifts 2, 5, and 9. In shift 7, the characters doing the action are mentioned by name in the previous spoken line, so a pronoun *they* suffices.

Four shifts contain information about a location. In shifts 5 and 11, the location is the same as it was in their previous similar situations, but with a focus to a specific place within that location. In shift 5, it is the box in front of which Moominpappa lies in the yard, and in shift 11, it is the window which the vines break in the attic. While the overall location is the same – the yard and the attic respectively – the specificity of the location warrants a more detailed description. Shift 9 occurs in a staircase, which is both a relatively new and a specific place. Although the characters have been inside the house and in other parts of the staircase before this point, the previous similar situation did not happen in a staircase. Therefore, the description is required. Shift 12 moves the action from a limited location, the roof of the Moomin house, to a general location, the Moominvalley. The more general location is then named, as while it contains the previous location, it does not focus solely on it anymore.

Shifts 2 and 7 do not mention a location. In the case of shift 2, the location can be inferable from the other information. The AD mentions a character, Mamelukki, and an object, a boat, which were both present in the previous similar situation and logically, to the Moomins at least, could not exist anywhere else than at the sea. The location in shift 7 is less obvious since the action, swinging on a vine, is pictured in a way that makes the background unnoticeable and less likely to be focused in the description. Timing seems to not be the issue in this case, as there is time to describe more in shift 7, which indicates that the describer may have decided not to include information about the location even when it could have fitted within the allowed time.

All four instances of *how* in this material appear in the shifts with more information, and it is the second-to-least common category overall. In shifts 2 and 7 it modifies an action. Shift 2 contains a description of the speed of the action, whereas shift 7 describes the perceived emotion of the characters as well as an additional visual clue of the action, *tails swishing*. Shift 5 describes the position the character is in (*on his stomach*). Shift 12 contains the only description on illumination in a scene with the evening *darkening*. The only description of *when*, the mention of *evening*, is in the last shift, which is also the last piece of AD in the episode.

All in all, the general tendency in the narrative surrounding these shifts is that they all contain something that is new, specific, or audibly ambiguous, which results in more information being relayed in the AD. Next, I will look into the differences between the situations which

require different amounts of information in the description and attempt to arrive at a conclusion as to why the direction seems to correlate with the amount of information in a shift.

4.3 Comparison of the two groups

As seen from the material, most shifts require a description of a location. This follows the classification by Marzà Ibañez, who places the *where*-information as the first and implicitly the most important one on her list. Location is also the only type of information that is sufficient alone in an utterance to convey that a shift in narrative is occurring. All shifts towards conversation contain *where*-information. It also seems that, when faced with a time restriction, the describer opted to describe a location rather than characters, as can be seen in shifts 6, 8, and 10.

What-information seems to correlate with the direction of the shift. This type of information is present in only three out of the seven shifts from story to conversation, whereas it is present in all five conversation-to-story -shifts. This tendency seems to represent the differences between the types of action that can occur in a story or during a conversation. A conversation held in spoken language requires people to speak out loud which restricts the types of action happening in a conversation to one type: audible action which does not require a description. On the other hand, what transpires in a story can be essentially anything a storyteller's mind can conjure, which does not restrict the level of audibility of an action. In this material, the action in a story often happens to be non-audible. Another reason for why *what*-information is so common in shifts towards a story might be the overall tendency of stories containing action, which is why they are told to others in the first place.

The *who*-category correlates with the shift direction as well. Only one story-to-conversation -shift contains information of this type, while all but one conversation-to-story -shifts have a *who* in them. Once again, this seems to be a result of the features of action in stories and in conversations. The conversations often contain the same cast of characters who listen to the stories without anyone standing out from the rest of the characters, save for the storyteller, so repeating *they* in the descriptions is redundant. On the other hand, when a storyteller tells a story, the focus is often on someone or something doing the action in the story, so the specific character or thing has to be described.

Information on *how* is balanced across the directions with two in each group, and I am not surprised by this result because of the nature of this type of information. Marzà Ibañez sets the *how*-category as the last in her listing, deeming it, at least subconsciously, the least important in the order of information. The information in this category modifies the action or the agent in a way that seems to be not exactly essential but rather something that is “nice to know”. This inclination seems to follow the discussion on “need to know versus nice to know” (Fryer 2016, 58), where a line has to be drawn in the importance of information to be described; what does the audience definitely need to know in order to understand a scene, and what information is something that is nice to know but not essential? Fryer places the *how*-information in the latter group as it describes visual details and descriptive elements without which the narration is still understandable and followable. This division makes the *how*-information a sort of a wild card that can appear whenever the visuals demand it and the timing allows for it to be included in the AD.

The final type of information in the material, the singular *when*, gives no indication on any correlation between the groups. It is placed in a shift from a story to conversation, but the shift overall is somewhat different from its group. It is a shift with four types of information, but its direction indicates that it should have less information. What, then, makes shift 12 so different narratively that it behaves irregularly in comparison to its group? To begin, it has a different visual method of conveying the shift. Usually, the shift is done visually by either the image freezing for a couple of seconds, a blackout, a blending, or a hard cut. This shift, however, has the overall location stay the same with only the characters blending out of the image. The visual combines a Moon rising above the Moomin house, and it seems to “swipe” the Moomins off the roof as it rises. Among no changes in the location, this shift has no change in the audio either, so for the non-seeing audience, the shift has to be indicated with some more information than the other shifts to conversation.

Other outliers from their respective groups are shifts 4 and 5 which, given their directions, seem to be in the opposite groups with regards to the amount of information in them. These shifts are related to the same story, with shift 4 moving to the story and shift 5 bringing the narrative back to the conversation. The story is Moomintroll’s own version of the fishing trip which he tells his father when asked to do so. However, the manner of storytelling is different from the other stories. In them, the content of the story is presented in a way that is indistinct from the narrative of the episode that is set in the realm of conversation. The characters talk to

each other as they normally would, and there is no outside narrator. In Moomintroll's story, however, the audio of the spoken lines remains in the realm of conversation while the visuals and the sound effects show and tell what happened in the taleworld. Moomintroll narrates the events of the story in past tense, acting as a kind of an audio describer himself. For this reason, shift 4 does not require as much description. The situation of the fishing trip is also known to the audience as it is the same as in the previous similar situation, further reducing the need for description. Then, shift 5 contains Moominpappa's evaluation in the form of him falling asleep mid-story out of apparent boredom. The situation contains a specific character and a specific location combined to an ambiguous sound, rendering it longer than usual shifts back to the conversation.

4.4 Division according to the interruption status

As I set out to study this episode further, I expected to see the interruption of a story affecting the AD in some way. As the research progressed, I noted that it was more insightful to inspect the material divided by the amount of information in each utterance. However, as it was my original meaning to study the different factors that affect the AD in this episode, I will briefly analyse the utterances divided by their interruption status as well.

Shifts 1, 2, 5, 6, 7, 8, and 9 are interrupted, while the rest are noninterrupted. The interrupted shifts have an average of 2.43 types of information in each utterance, while the noninterrupted shifts have an average of 2.2 types of information. *Where* divides evenly with five instances in each group. *What* is more often described in interrupted shifts with five instances compared to the three in noninterrupted shifts, but they take up similar percentages of the overall shifts in the groups – 71% and 60% respectively. There is more dispersion in the rest of the categories with *who* present in four interrupted shifts but only in one noninterrupted shift, *how* present in three interrupted shifts and one noninterrupted shift, and the singular *when* in a noninterrupted shift.

Some of these differences could be explained by the definition of an interrupted shift. Earlier, I defined an interrupted shift as a shift that lacks one or more frames from the order of Young's theory. All of the interrupted shifts towards a conversation are interrupted by an evaluation from a listener, which skips other frames that should precede an evaluation, whereas the interrupted shifts towards a story lack frames from either the very beginning of

the frame cycle or very close to the story itself. This tendency could be because the story that follows the interruption suffices with the same preface as earlier, since the speaker still has their storytelling turn and does not have to present their attempt to tell the story another time. Another reason could be that the story continues where the interrupted story ended, so the frames very close to the story are not required as they have already been told. The lack of definitive types of interruption means that it is difficult to draw conclusions for interruption's effect to AD.

5 Conclusion and discussion

Let us now return to the research questions and see how the analysis answers them.

1. How does the AD support the narrative shift from the story to the conversation and back?

Information about a location seems to be the most important type of information that has to be conveyed in the AD. In this type of narrative, it is also logical that a location is most often described, since it is an obvious visual clue about a shift in narrative but hardly ever inferable from audio only. Location also seems to be preferred over other types of information when time is scarce and only a short description fits between the other sounds in the episode.

It seems that the direction of a shift correlates somewhat with the amount of information in an utterance. The AD contains less information – 2 types per utterance in average – when the narrative returns to the conversation from a story. Conversely, the AD contains more information – 2.8 types per utterance in average – in the shifts towards the story. There seem to be two main reasons for these tendencies. Firstly, the action in a conversation is inherently audible, which reduces the need for a description in this direction. On the other hand, a new story often contains a new action with ambiguous audio, which in turn requires a description. Secondly, the situations where the conversations occur often have the entire cast of characters present, so there is less requirement to mention all of them separately. In contrast, the stories often include a specific character or characters doing an action, which results in them being named more often.

2. What differences and similarities can be drawn from the material based on the factors present in the shifts, such as the direction of the shift or its interruption status?

Before analysing the material further, I had assumed that the interruption status of a shift affects the AD somewhat. Surprisingly, this turned out to not be the case, and instead, the factor that had more effect on the AD was the direction of the shift. As I was able to see the visuals of the episodes, I thought that the freeze images in the interrupted shifts would have been a bigger issue, requiring a description. Instead, none of the different types of cutting between the story and conversation were mentioned in the AD at all. At first, I assumed this was a fault on the describer's part; surely a description is required in a distinctive visual like a plant beast being about to attack the main heroes, and the image then freezing for almost ten seconds in the point of highest suspension with a singular character moving in the image with the others frozen.

After some thought and familiarizing myself with AD theory, I realised that leaving the striking visual effects undescribed could have been a conscious decision on the describer's part, and a reasonable one at that. When audio describing to children, one has to keep in mind that the target audience is likely to have less or no memories of having been able to see (Fryer 2016, 114) when compared to an audience of adults, who are more likely to have developed visual impairments later in the life. This means that a blind or visually impaired child might not have internalised how a video works and thus, has no idea what an image freezing looks like or means in the context of a film or a TV-show. Therefore, this type of information that is conveyed through purely visual details might not be as relevant to a visually impaired audience as they can be to a seeing one. Also, in this episode at least, the striking effect is accompanied by music that behaves similarly to the visuals by stopping as the image stops, which can also relay the information about something unusual to the audience.

The results of this study are not generalizable to a larger amount of data because of its small sample size. It does give some indication on what type of information is essential in a specific narrative structure. However, repeating the same study with different material could prove to be difficult and might give different results. This study requires two phenomena to coincide: a film or a TV-show must have a structure of a framed narrative, and such material must also be audio described. In Finland, the latter is still regrettably rare, and the former is incidental and up to the writers and creators of film and TV to decide.

More fruitful avenues of future research include research on children's audio described material in general. While Näkövammaisten kulttuuripalvelu ry has guidelines to AD (2012)

in Finnish, they are very broad and do not discuss how one should describe to children. It would be interesting to compare a piece of children's described media to one that is geared for an adult audience to see if the lack of a proper guideline for describing to children in Finnish has affected the scripts.

In Finnish media, AD is mostly done by a small and limited amount of people, as can be gleaned from the list of Finnish audio described media (Luettelo Suomessa kuvailutulkatuista elokuvista ja TV-ohjelmista, n.d.). Much of Finnish AD production might be influenced by the fact that there is little research, only few sets of guidelines, and a limited cast of describers. Notably, describers Anu Aaltonen and Heikki Ekola are present in over a half of the productions listed on the website (*ibid.*), which might result in their personal working styles being overrepresented in the Finnish AD content. It could also be that one describer's tendencies do not directly translate to another describer's style. A comparative study on different describers' styles would give insight on how uniform or scattered the prevailing state of AD is in Finland.

6 References

6.1 Research data

”Jakso 8: Muumipapan kalajuttu.” *Muumilaakso (kuvailutulkattu)*, directed by Steve Box (Finland/Great Britain: Gutsy Animations, 2019), Video. Audio described by Anu Aaltonen, Hannele Antikainen and Carita Lehtniemi. Accessed January 29, 2020. <https://areena.yle.fi/1-4664041>

6.2 Other references

Heinmaa, Heidi. “Kuvailutulkkausohjeiden parhaimmat käytänteet ja ohjeiden sovellettavuus.” Master’s thesis, University of Helsinki, 2017.

Hirvonen, Maija. “Multimodal Representation and Intermodal Similarity – Cues of Space in the Audio Description of Film.” Doctoral dissertation, University of Helsinki, 2014.

Fryer, Louise. *An Introduction to Audio Description: A Practical Guide*. Oxon: Routledge, 2016.

Marzà Ibañez, Anna. ”Evaluation criteria and film narrative. A frame to teaching relevance in audio description.” *Perspectives: Studies in Translatology* 18, no. 3 (2010): 143–153. <https://doi-org.libproxy.tuni.fi/10.1080/0907676X.2010.485682>.

Nuutinen, Maija. ”Näkövammaisten kuvailutulkkauskokemukset.” Master’s thesis, University of Eastern Finland, 2014.

Näkövammaisten kulttuuripalvelu ry. ”Kuvailutulkkaustoimikunnan ohjeet kuvailutulkille 2012.” Accessed January 29, 2020. <http://www.kulttuuripalvelu.fi/kuvailutulkkaus/ohjeita>

Näkövammaisten kulttuuripalvelu ry. ”Luettelo Suomessa kuvailutulkatuista elokuvista ja TV-ohjelmista.” Accessed April 20, 2020. <http://www.kulttuuripalvelu.fi/kuvailutulkkaus/kttietopankki>

Palomo López, Alicia. ”The benefits of audio description for blind children.” In *New Insights into Audiovisual Translation and Media Accessibility : Media for All 2*, edited by Jorge Díaz Cintas, Anna Matamala and Josélia Neves, 213–26. Amsterdam: Rodopi, 2010.

Pesonen, Laura. ”Nur eine dunkle Vorstellung von den Bildern? Bildende Kunst mit Audiodeskription, Struktur und Inhalt der Audioskripte.” Master’s thesis, University of Eastern Finland, 2015.

Piety, Philip. ”The Language System of Audio Description: An Investigation as a Discursive Process.” *Journal of Visual Impairment and Blindness* 98, no 8 (2004): 453–469.

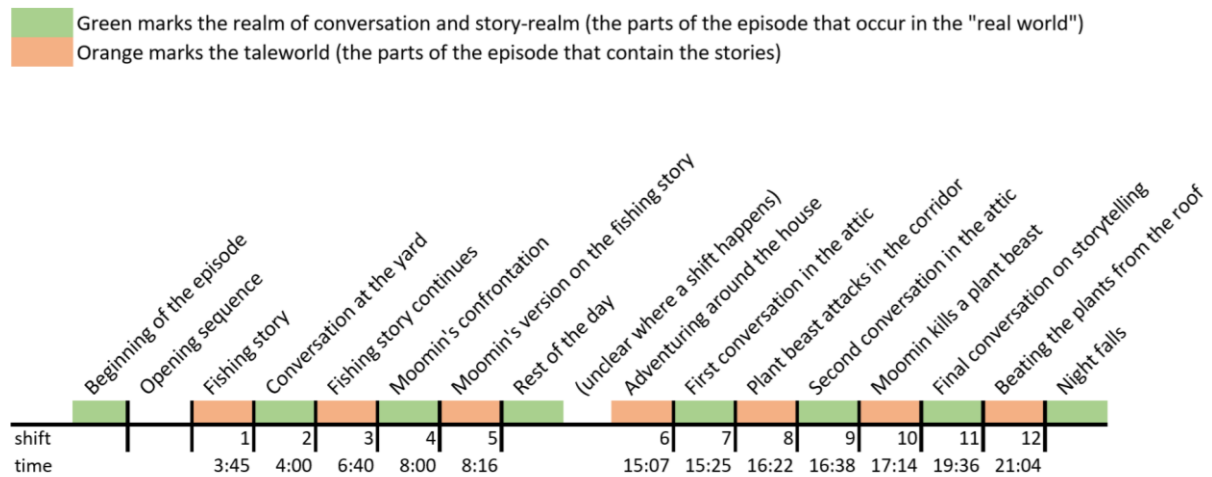
Reiman, Outi. ”Kääntyykö kuvailutulkkaus? The Big Short -elokuvan kuvailutulkkeen suomennettavuus.” Master’s thesis, Tampere University, 2017.

Taylor, Christopher. ”The multimodal approach in audiovisual translation.” *Target* 28, no. 2 (2016): 222–236. <https://doi.org/10.1075/target.28.2.04tay>.

- Vaittinen, Emma. "Depictions of men and women in audio descriptions: Character physicality in ADs of action films." Master's thesis, University of Helsinki, 2018.
- Valli, Heli. "Kuvailu ja haptiisit näkövammaisten opiskelijoiden kotitalousopetuksessa." Master's thesis, University of Helsinki, 2017.
- Young, Katharine. "Frame theory." In *Routledge encyclopedia of narrative theory*, edited by David Herman, Manfred Jahn and Marie-Laure Ryan, 185–186. London: Taylor & Francis, 2010.

7 Appendices

Appendix 1. Timeline of the episode with the shifts, time stamps, and colouring to indicate realms of conversation and taleworlds



Appendix 2. Utterances and their translations divided into the categories

Shift	Utterance	Translation [all my translations]	Category
1	Istuvat	(They) sit	what
	kotipihalla	at the home yard	where
2	Mamelukki	Mamelukki	who
	alkaa vetää venettä	begins to pull the boat	what
	kiihtyvällä vauhdilla	at an increasing speed	how
3	Kotipihalla	At the home yard	where
4	Ovat merellä	(They) are . . . at the sea	where
	onkimassa	fishing	what
5	Pappa	Pappa	who
	kuorsaa	snores	what
	vatsallaan	on his stomach	how
	laatikon edessä	in front of the box	where
6	Talon ullakolla	In the attic of the house	where

7	Onnellisina	Happily	how
	he	they	who
	liitävät liaanilla	glide on a vien	what
	hännät viuhuen	tails swishing	how
8	Ullakolla	In the attic	where
9	Pappa, Mamma ja Niiskuneiti	Pappa, Mamma, and Snorkmaiden	who
	perääntyvät	back down	what
	portaikossa	in a staircase	where
	kasvipedon	(as) a plant beast	who
	hyökätessä	attacks	what
	heitä kohti	towards them	where
10	Ullakolla	In the attic	where
11	Köynnökset	Vines	who
	tunkeutuvat	push	what
	ikkunoiden läpi	through the windows	where
12	Ilta	The evening	when
	pimenee	darkens	how
	Muumilaaksossa	in Moominvalley	where
	ja kuu nousee taivaalle	and the Moon rises on the sky	what