From Unisemiotic to Polysemiotic Narratives:

Translating across semiotic systems

Karoliina Louhema

Supervisors: Jordan Zlatev, Maria Graziano, Joost van de Weijer

MA Program in Languages and Linguistics
Specialisation: Cognitive Semiotics
MA Thesis, 30 ECTS
June 2018

Abstract

Human communication is both *polysemiotic* and *multimodal*; it is comprised of ensembles of representations from different *semiotic systems* in different *sensory modalities*. These semiotic systems, such as language, gestures, and pictures consist of signs and relations between signs, with system-specific affordances (Kendon, 2004; Zlatev, 2009; Sonesson, 2014). These systems all have their unique storytelling potentials, and how meaning is conveyed and what is communicated may differ across systems and modalities (Kress, 2010; Green, 2014). This thesis investigates the influence of the source semiotic system on the way the story content is transmitted from unisemiotic and unimodal narratives to polysemiotic and multimodal narratives. The source systems were language in the speech narrative mode (SNM) and a sequence of pictures in the picture narrative mode (PNM). An experimental method was used to see whether there are differences in the polysemiotic narratives – that is, in the retellings through speech and gestures – that could be connected to the semiotic system in which the narratives initially were perceived. The participants were 38 native Finnish speakers, and the experiment was carried out in Finnish. Due to the system-specific differences, it was hypothesised that pictorial iconicity would be reflected as a higher number of iconic gestures – especially first person enactments – and ideophones in PNM, and SNM would result in greater narrative coherence reflected, for example, in the more diverse use of connective devices and a higher number of plot elements.

Contrary to expected, more iconic gestures were found in the narratives translated from the SNM condition. However, in line with the hypotheses, first person enactments were more frequent in the narratives of the participants who had experienced the story through pictures. Also contrary to expected, the difference in the use of connective devices between the conditions was not remarkable. However, together with some additional differences between the groups that had not been anticipated, the results indicate that a story given in different semiotic systems may indeed lead to different polysemiotic narratives.

Keywords: bodily mimesis, cognitive semiotics, co-speech gestures, embodied narratives, experientiality, iconicity, ideophones, intersemiotic translation, language, mimetic schema, multimodality, narratives, phenomenology, pictorial semiotics, pictures, polysemiotic communication, semiotics, semiotic systems, sensory modalities, signs, sound symbolism

Acknowledgements

First and foremost, I would like to express my deep gratitude to my main supervisor Jordan Zlatev, who has patiently been there guiding me throughout this long journey. Since I had worked with the frog story already in my BA thesis, examining language development of little L1 English speakers, Jordan passed me this challenge, and it has taught me a lot along the way. During this excursion I have gone through numerous sources, and been lucky to get familiar with so many aspects involved in cognitive semiotics. To understand the bigger picture, I wanted to know more about everything. When I got lost in the theory jungle, Jordan gently ushered me back on track, and he was always available to discuss and brainstorm, share the excitement and develop ideas, share the knowledge, and plan ahead. Jordan always has an answer to everything, and he's always on the same page!

I would also like to thank warmly my second supervisor Maria Graziano for your expertise in gesture analysis, your involvement, sparkly presence – and patience! My third supervisor, Joost van de Weijer, thank you for joining in, being available even at short notice – making time, and being excited with us! Thank you for your contribution. I am also grateful for the inspiring learning environment at Lund University, and the encouragement of the teachers. Furthermore, I am happy to have met all my course mates at SOL. We were a small bunch, and I enjoyed our discussions and loved the group support. I'm looking forward to seeing where your careers take you!

Thank you Anu Vastenius and Lari-Valtteri Suhonen for piloting the experiment with me. Big thank you Anna Kiuttu for lending your voice to the frog story! It is your voice that I now hear when I think about the story. Also, thank you everyone who took part in the experiment, especially the 38 participants without whom this experiment would never have succeeded. Spending hours and hours working with the videos was fun because of you. Your narratives reflected your personalities, and made me chuckle – so much! I also received help from many of my friends assisting in the test situation, and also outside the experiment helping, for instance, with translation judgements. Thank you! Special thank you to Kaius Louhema for being my research assistant on many occasions, helping me with this and that, and being ever so supportive! You played a big part.

I also thank my mother. I thank all my friends in Copenhagen-Malmö-Lund area, for always being keen to catch up and going to the pub quiz and offering me a couch to stay when I'm around and what not! I cherish our friendships. Thank you Minna Tanskanen for always inviting me over and for the years we have known each other. You've become a dear friend.

Naturally, I also want to thank my two furry sons, C & C, for taking me out for walks to reflect and experience the real-world wonders in the real-world forests and breath.

TABLE OF CONTENTS

LIST OF FIGURES I

LIST OF TABLES II

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 THEORETICAL FRAMEWORK 4

2.1. Introduction 4
2.2. Cognitive semiotics 4
2.3. Semiotic systems and polysemiotic communication 8
2.3.1. Language 8
2.3.2. Gestures 12
2.3.3. Pictures 15
2.3.4 Polysemiotic communication 18
2.4. Stories, narratives, and narrativity 20
2.4.1. Concepts 21
2.4.2. Narrativity and narrative structure 21
2.4.3. Narrative modes 23
2.4.4. Embodied narratives 28
2.5. Translating between semiotic systems 29
2.6. Summary 31
2.7. Aims and general hypotheses 31

CHAPTER 3 METHOD 33

3.1. Design 33
3.1.1. Materials 33
3.1.2. Participants 33
3.1.3. Ethical considerations 34
3.1.4. Setup 34
3.1.5. Procedure 34
3.2. Speech annotation and coding 35
3.2.1. Clauses 35
3.2.2. Plot elements 36
3.2.3. Connecting devices 37
3.2.4. Story characters 38
3.2.5. Ideophones 38
3.2.6. Gestures 39
3.3. Specific hypotheses 39
3.4. Summary 40

CHAPTER 4 RESULTS AND DISCUSSION 41

4.1. Introduction 41
4.2. Speech analysis 42
4.2.1. H1 – Plot elements 42
4.2.2. H2 – Connective devices 43
4.2.3. H3 – Story participants 44

4.2.4. H4 – Formulaic expressions from the genre of storytelling *44*
4.2.5. H5 – Ideophones *45*
4.2.6. H6 – Ideophone modality *46*
4.3. Gesture analysis *47*
4.3.1 H7 – Iconic gestures *47*
4.3.2. H8 – Enacting gestures *48*
4.4. Discussion *48*
4.5. Summary *58*

CHAPTER 5 CONCLUSIONS 60

REFERENCES 63

APPENDIX A FROG, WHERE ARE YOU? (ADAPTED FROM MAYER, 1969) 70

APPENDIX B FROG, WHERE ARE YOU? TRANSCRIPT (ENGLISH) 73

APPENDIX C FROG, WHERE ARE YOU? TRANSCRIPT (FINNISH) 75

APPENDIX D FEEDBACK FORM 77

LIST OF FIGURES

Figure 1. The conceptual-empirical spiral 5

Figure 2. The triadic relationship between expression, content, and interpreter 6

Figure 3. Primary iconicity 7

Figure 4. Droodle 7

Figure 5. The upper and lower limits of gesture 12

Figure 6. The level of iconicity from enactment to symbolic gestures 14

Figure 7. Pictorial consciousness 17

Figure 8. Proportions of narrative clauses 43

Figure 9. Proportions of metanarrative clauses, extranarrative clauses, and narrative clauses 43

Figure 10. Proportions of connective devices in categories of conjunctive relation 44

Figure 11. Proportional ratio of main and peripheral characters 45

Figure 12. Proportion of formulaic story expressions 45

Figure 13. Ideophone rate per narrative clause 46

Figure 14. Proportions of ideophone modality 46

Figure 15. Gesture rate per clause and per word 47

Figure 16. Proportional division of gestures in functions 47

Figure 17. Proportion of enacting gestures counted from iconic gestures 48

Figure 18. Picture 10 in the picture source narrative that presents the gopher biting the boy's nose while the dog is harassing the bees 52

Figure 19. Picture 16 in the picture source narrative that corresponds with line 19 in the English transcript 53

Figure 20. Proportion of first person direct speech and third person indirect speech perspective 54

LIST OF TABLES

Table 1. Ideophones and phonestemes 10

Table 2. List of Finnish clause initial connective devices divided in categories of conjunctive relation 37

Table 3. Examples for ideophone inclusion 39

Table 4. Statistical difference between PNM and SNM in H1-H8 41

Table 5. Overview of the results related to H1-H8 42

Table 6. Differences in the proportions of iconic gestures 47

Table 7. A list of interjections in PNM 57

CHAPTER 1 INTRODUCTION

> Telling stories is hard. First you need something to tell – a genuine *story*. You must be acquainted with a happening, or series of events, about which you can make a point. […] Even before you start, you need an overall plan – where to begin, what events and evaluations to include or emphasize, and where to end. In the actual telling, you need to *formulate* utterances one by one. For each utterance, you must select the right *words and gestures from the repertoire available* to you – your *language* and *culture* – and you must do so in a timely fashion. Your selection must be both true to the events in the story and understandable to your audience.
>
> Herbert Clark, 2004 (my emphasis)

I start by quoting Clark (2004: 457), because ultimately this thesis is a study of *narratives* – of the telling (and understanding) of a *story*[1] – and this quote nicely captures the features that should be acknowledged from the start. Not only does it capture the significance of *language*, but also that of expressive bodily movements – *gestures,* a semiotic system of their own, though closely connected with speech. It also captures the role of *cognition* in this process. *Narrative thought* is a human quality that we use to make sense of the world around us and it is possible that language initially evolved precisely for the sake of narrative (Donald, 1991). It is a view held by many that "[m]apping words onto worlds is a fundamental – perhaps the fundamental – requirement for narrative sense making" (Hühn et al. 2009: 37). The idea of language evolving to enable the formulation of complex narratives could help explain the existence of structural devices for narrative coherence in all languages (Tomlin & Myachykov, 2015). These allow the easy tracking of discourse referents and other discourse cues important for mapping causal and other relations between events.

The cognitive-semiotic factors at play allow the perception of narratives through the different senses – or *modalities* (i.e. visual, auditory, kinetic, or haptic) – which is deployed effectively in human cultures by creating meanings with cultural artefacts, such as, paintings, music, plays, books, and dance. Indeed, there are many ways a story could be presented apart from language: in pictorial form, film, pantomime, etc., and the semiotic systems that allow these different ways of representing a story are of special interest in this study, specifically those of pictures, language, and gestures. When used for producing narratives the different semiotic systems give rise to specific *narrative modes.*

---

[1] Following the lead of Berman & Slobin (1994), and Clark (2004), this thesis makes a distinction between a *story* and a particular *telling* of that story by respectively calling them the *story* and the *narrative,* thus differentiating the two notions.

Kress (2010: 22) stresses the importance of research in *multimodality*, in his sense involving different "modes, each chosen from rhetorical aspects for its communicational potentials"[2]. For example, a picture in a road sign or advertisement may work better than a lengthy text, and a hand gesture in a discourse can emphasise or hint to what a linguistic message leaves out. An alternative, and less-ambiguous way to state this, is that different semiotic systems have their unique storytelling potentials and can be used to deliver a story, but how the meaning is conveyed and what is communicated may differ across systems and sensory modalities. This could be explained by the system specific affordances[3], which affect how a story is narrated and experienced (Hühn et al. 2009; Kress, 2010; Green, 2014). When comparing storytelling in language and pictures, the latter may have less cohesion compared to the former, but be more expressive and detailed, for example, in the setting, alignment, and spatial organisation of the elements (Hühn et al. 2009). Moreover, multimodality clearly presents itself in face-to-face communication in the form of *embodied narratives* that are shaped by "the unfolding sequence of embodied and embedded perceptions of an individual" (Menary, 2008: 76). This kind of multimodal interaction combines a number of semiotic systems and resources, such as *speech*, *manual gestures*, *head* and *body movements*, *facial expressions*, and *pictures,* which are all central for human communication (Streeck, Goodwin, & LeBaron, 2011). Recently, methods from *phenomenology*[4] have been used in attempts to explore how "meaning is intrinsically layered in human experience" (Zlatev & Blomberg, 2016:185) and re-enacted in our performances. Such layering of meaning occurs in both ontogeny (e.g. Piaget, 1962) and phylogeny (e.g. Donald, 1991).

This thesis addresses questions that aim to contribute to issues researched in *cognitive semiotics*, concerning how the *narrative content* (i.e. story) may be influenced by the *narrative mode*, as well as bring new insights into the interaction between the semiotic systems of language and gestures (Zlatev, 2015). Gesture research has still been very slim in connection to Finnish language, with the focus predominantly on communicative difficulties (Jääskeläinen, 2009;

---

[2] Kress (2010) sees the selection of "modes" in human communication more broadly than many other researchers, for he treats units, such as, *colour, prosody, touch, music, gesture*, and *speech* all equally as such "modes" for creating meaning, but according to Green (2014: 10), this approach is problematic, because this "leads to an abundance of modes that are difficult to compare".

[3] The term *affordance* should not be understood here in the sense of Gibson (1979), but rather as potentiality for expressiveness. That is, when comparing how different semiotic systems may be used to communicate information, one may be more restricted in some sense than another while conveying other type of information more directly than the other. Thus, the two have different affordances (and restrictions) – i.e. expressive potentials for meaning making.

[4] Phenomenology is a philosophical tradition used to study human experience. One of its main ideas is that our thinking and how we perceive things around us is moulded by our experiences in the perceptual world that we are part of (Sokolowski, 2000).

Haddington & Kääntä, 2011). Thus, empirical evidence based on healthy native speakers will also make a contribution to Finnish language research. At the core of this thesis is an experimental study with intersemiotic emphasis, which broadens and deepens previous research on narratives by combining traditional methods with novel ones. Comparing verbal narratives to a sequence of static pictures as a narrative mode could bring valuable insights to narratology (Sonesson, 2015a), as pictorial material and polysemiotic narratives have scarcely been under experimental investigation in narratological research as a result of the dominant focus on language (Hühn et al. 2009; Kress, 2010; Ranta, 2011; Yiheng Zhao, 2015). By comparing two groups of participants exposed to a story in two different narrative modes (i.e. narratives delivered in two different semiotic systems: language and a sequence of pictures), it is possible to test whether they would result in mode-specific differences when later translated into a polysemiotic narrative (i.e. retelling of the story in speech and gestures). The questions that this thesis asks are thus:

- What happens when the same story expressed in either language or pictures is translated into an embodied narrative?
- Would hearing the story give more coherent embodied narratives in respect to organisation and the development of the plot?
- Would seeing the story give rise to more perceptually detailed narratives, including iconic gestures and ideophones (clearly non-arbitrary, though still conventional expressions)?

This thesis is divided in five chapters. The theoretical background is presented in Chapter 2, which builds up the framework of the study. The aims of the study and research questions are taken up once again at the end of the chapter with this framework in place, and the methodology and design of the experiment together with specific hypotheses are explained in Chapter 3. Chapter 4 presents the results and discussion, and finally, conclusions are given in Chapter 5.

CHAPTER 2 THEORETICAL FRAMEWORK

2.1. INTRODUCTION

This chapter presents the theoretical framework of this thesis by first introducing the field of *cognitive semiotics*, which builds up the core background for the study. It then moves on to introduce the most relevant concepts for the present research, specifically, the semiotic systems of language, pictures, and gestures, and the interaction of different systems in *polysemiotic communication*. Furthermore, the chapter leads us to the study of narratives, and describes the narrative modes compared in the study. Also, a detailed description is offered of how the concept of *translation* is to be understood in the study, and how it is relevant for the topic. Finally, a chapter summary together with a more detailed specification of the research questions leading to general hypotheses round up the chapter.

2.2. COGNITIVE SEMIOTICS

Cognitive semiotics has emerged over the past decades to accommodate the growing need across a variety of research fields to help understand mind, life, language, and society better (Zlatev, 2012; Sonesson, 2015a; Sonesson & Lenninger, 2015). This field is dedicated to the "transdisciplinary study of meaning", combining knowledge and methods from semiotics, linguistics, philosophy, cognitive science, and developmental and comparative psychology (Zlatev, 2012: 2). Some topics that have been in focus are bio-cultural evolution, semiotic development in ontogeny, "the embodied mind", and multimodal communication (Zlatev, 2015a). Researchers in cognitive semiotics aim to integrate theoretical and empirical research favouring the use of methodological triangulation, including the use of a combination of 1st person (e.g. intuition), 2nd person (e.g. empathy) and 3rd person (e.g. experimentation) methods (Zlatev, 2009). In addition, cognitive semiotics is influenced by *phenomenology* (especially the work of Edmund Husserl, and Maurice Merleau-Ponty) to help understand the complex relationship between bodily experience, sociality, and language. Phenomenological analysis is a method that can provide a careful description of how things and acts of meaning making present themselves to us in experience (Sokolowski, 2000; Zlatev, 2009). One could say that:

> the basic idea is to depart from experience itself, and to provide descriptions of the phenomena of the world, including ourselves and others, as true to experience as possible – rather than constructing metaphysical doctrines, following formal procedures, or postulating invisible-to-consciousness causal mechanisms that would somehow "produce" experience. (Zlatev, 2012:15)

One of the basic insights of phenomenology is that we live in the *lifeworld* (in German: *Lebenswelt*, a term coined by Edmund Husserl) that we share with other beings, perceiving and acting in this

already meaningful world that is both personal and *intersubjective.*[5] That is, we are shaped by our experiences in the lifeworld through our living and lived bodies. According to Husserl's notion of time consciousness, all present moments contain references to past moments (retentions), which themselves have references to yet earlier moments resulting in a type of accumulation of meaning – in "the layering of meaning over meaning in time" (Sonesson, 2015a: 37). At the same time, the present moment is ripe with future possibilities, protentions. On a larger temporal scale, Zlatev and Blomberg (2016) discuss the notion of *sedimentation* (i.e. the process of meaning being layered in the lifespan, and across generations) to demonstrate how linguistic meaning is from the start grounded in both our bodily experiences and sociocultural practices. Zahavi (2003: 104) argues that "sociality presupposes a certain intersubjectivity of the body", and thus the term "shared mind" is better suited to be used in connection with intersubjectivity than the notion of "theory of mind" (Zlatev, 2008b). The notion of intersubjectivity is important to acknowledge when studying communication – especially in tête-à-tête situation, which involves "the sharing of affective, perceptual and reflective experiences between two or more subjects" (ibid: 2015).

One of the central characteristics of cognitive semiotics is that it underlines the importance of delivering new empirical research that not only contributes in the work of clarifying concepts like *sign, language, representation, meaning*, and *intersubjectivity*, but also applies these in empirical studies in a way that creates new insights (Zlatev, 2015a). This type of *conceptual-empirical spiral* (see Figure 1) that Zlatev (2015a) calls this characteristic of cognitive semiotics is essential also for the present thesis.

Figure 1. The conceptual-empirical spiral (adapted from Zlatev, 2015a: 1058)

One of the principal concepts in cognitive semiotic research is that of the *sign*. Sonesson (2010) defines a sign as something that contains at least the following two parts: *expression* (representamen) and *content* (object). Between the expression and content there is a double

[5] The phenomenological notion of intersubjectivity emphasises that "the subjectivity that is related to the world only gains its full relation to itself, and to the world, in relation to the other" (Zahavi, 2001: 166). Furthermore, Zahavi (2003: 105) argues, following Husserl, that "we cannot perceive physical objects without having an accompanying bodily self-awareness [...]. But the reverse ultimately holds true as well: The Body only appears to itself when it relates to something else – or to itself as the Other."

asymmetry, because the former is more directly perceived than the latter, whereas the latter is more in focus than the former. Simply put, the expression is something that can bring something in awareness other than itself. Therefore, the expression and the content must be clearly differentiable from one another as well as linked (Zlatev, 2009). Moreover, the relationship between expression and content, and the person interpreting their relation makes this a threefold relation (see Figure 2), for something cannot be a sign unless the connection and difference between expression and content can be interpreted by consciousness (Zlatev, 2008a). For something to be seen as a sign, there needs to be a pre-existent relation between the expression and content (although the relation itself does not assume the sign function) (Sonesson, 2015b). The relation, termed *ground*, can be of three kinds: *iconic* – based on similarity, *indexical* – based on association (e.g. spatiotemporal or a part-whole relation), and *symbolic* – based on convention (Sonesson, 2010). This division follows the "trichotomy" of the well-known American philosopher Charles Peirce, but any concrete sign can be based on a combination of these different grounds (i.e. iconicity, indexicality, and conventionality), and it is the predominant ground of the sign that determines its type – an iconic sign, indexical sign, or a symbol (Jakobson, 1965).

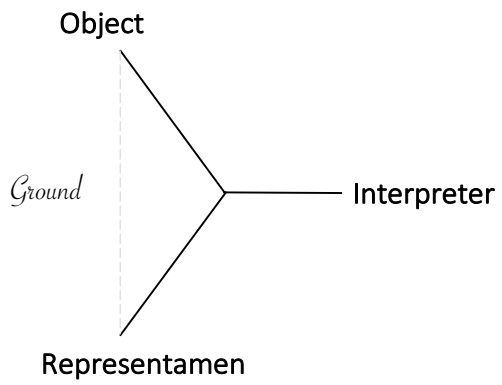

Figure 2. The triadic relationship between *representamen* (expression), *object* (the content that the expression stands for), and *the interpreter* (adapted from Ahlner & Zlatev, 2010: 314).

Two entities can be said to share an iconic ground when they both have – independently of one another – a set of properties that are "identical or similar when considered from a particular point of view" (Sonesson, 2010: 29). The two items thus resemble one another, like an apple (the object) and a picture of it (the representamen) (Figure 3). Iconic signs can be divided further in terms of (combinations of) *primary* and *secondary* iconicity (ibid). In the former the perception of similarity between representamen (expression) and object (content) is at least one of the reasons experiencing this as a sign, as in Figure 3. In the latter, it is only once expression is known to stand for something

(i.e. content), the resemblance between the two things becomes in focus. Such can be the situation if the expression is too ambiguous or if it may not be seen as a sign but as an object in its own right. Secondary iconicity is seen in so-called *droodles*, shown in Figure 4.

Figure 3. Primary iconicity. Viewing from the left we can see a photograph of an apple (object), and on the right a realistic picture of an apple (representamen).

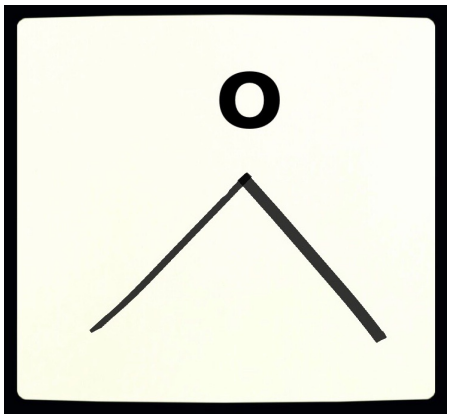

Figure 4. Droodle – an instance of secondary iconicity, where, until given away by context or prior knowledge, the image may not make much sense. Perhaps the droodle depicts the sun over a tent or a mountain, or a simplistic human figure.

The second of the three grounds – indexicality – is based on time/space *contiguity* or *factorality* (part-whole relation) between two entities: as smoke is associated to fire, or a head is part of the body. Just as iconicity, indexicality may be perceived, "but is not a sign until it enters the sign relation" (ibid: 30). Other examples of indexical signs could be a knock on the door as an index of someone at the door (Sonesson, 2015b) or the classical performative index, the pointing gesture (Andrén, 2010).

Lastly, in symbolic signs the predominant relation between expression and content is conventional. That is, the relation is based on the agreement between the sign users. Sonesson (2010) mentions traffic regulations as an example of symbolic signs, and Zlatev (2009) emblematic gestures (e.g. *OK* -gesture), words, and grammatical constructions.

2.3. SEMIOTIC SYSTEMS AND POLYSEMIOTIC COMMUNICATION

Semiotic systems like language, gestures and pictures consist of signs, relations between signs, and systemic-specific affordances (i.e. potentials for meaning making). In addition, Green (2014), suggests that language could be viewed as a superordinate semiotic system, which can be realised as spoken language, written language, and signed language, as each of these has specific expressive potentials, and involves different sensory modalities. However, in this subsection the focus is in language in general, and the semiotic subsystem of speech is taken separately later in connection with narrative modes. After introducing the semiotic systems of language, gestures, and pictures, the notion of polysemiotic communication – that is, the combination of separate semiotic systems – is introduced.

2.3.1. LANGUAGE

Language is predominantly (for there is also iconicity and indexicality involved) built of symbolic signs that are spoken, written or signed (Jakobson, 1965). Thus, language can be defined as "a conventional-normative semiotic system for communication and thought" (Zlatev, 2008a: 37):

> It is *conventional*, in the sense that it is based on an implicit or explicit agreement among its users. Even when implicit, these conventions are *normative*, in the sense of prescribing criteria for correct use. These conventions/norms are *semiotic*, since they take the form of signs […], organized in a system. And finally this system is used not only for *communication* (through speech, signing, or writing) but also in *thinking*, after being internalized by children.

This approach has its roots in early 20th century Swiss linguist and semiotician Ferdinand de Saussure's view of language as a social institution in contrast to, for example, *generative grammar*[6] (Chomsky, 1975) or *cognitivism*[7] (Lakoff & Johnson, 1999) with their emphasis on innateness and unconscious processes in the production of language. But the structuralism of Saussure focuses only on the anonymous and static *langue*. Language, however, is a more multi-faceted phenomenon than this. An important distinction was made by Coseriu (1985) between language as activity *(energeia)* – that is, the act of speaking, or language use in general; as knowledge *(dynamis)* – the knowledge of the vocabulary and grammar, etc. that enables an individual to understand and produce certain language; and as the product (*ergon)* of language activity, such as a spoken or written text (see Zlatev, 2011). The cognitivist and generative approaches to language mentioned

[6] Generative grammar (e.g. Chomsky, 1957) sees language deriving from a "mental organ" rather than being social and normative system, and focuses on syntax rather than meaning.

[7] Lakoff and Johnson (1999) argue that linguistic knowledge and the cognitive processes involved are not conscious processes, but rather derive from the "cognitive unconscious".

earlier are mainly focused on viewing language from the perspective of linguistic knowledge. However, the three perspectives are equally important and "necessarily co-imply each other" (Zlatev, 2011: 129). Not only does linguistic knowledge underlie language activity (and is a prerequisite for it) and the linguistic product (e.g. speech or text), but the contrary is also true, for the reason that language is based on social interaction and thus changes in use over time.

Being conventional does not imply being "arbitrary" (Ahlner & Zlatev, 2010). Iconicity on the level of single linguistic signs is reflected in *sound symbolism*, of which the most familiar to a layman would be *onomatopoeia*. Iconicity in onomatopoetic words is conveyed through "a degree of resemblance between the words' sound patterns and the sounds produced by certain animals or inanimate objects" (ibid: 306). Such resemblance is evident in words like *buzz*, *click*, *meow, snap*, and *splash* – with rather obvious similarity in sound between expression and content. But as linguistic signs are symbols, such similarity combines with conventionality. Denoted actions (1-4) show a typical example – the rooster's cry – in four different languages, which demonstrates both the similarities they all share with the actual cry sound (iconicity), and differences in English (1), Spanish (2), Finnish (3), and Dutch (4):

(1) *Cock-a-doodle-doo (EN)*
/ˈkɒk ə ˈduːdl̩ ˈduː/

(2) *Quiquiriquí (ES)*
/kikiriˈki/

(3) *Kukkokiekuu (Fin)*
/kukːo kiekuː/

(4) *Kukeleku (NL)*
/kykələky/

Iconicity in language does not necessarily have to involve the same sensory modality in both expression and its content, as in onomatopoetic expressions (i.e. audition), but a variety of different "sensory scenes" may be imitated by phonetic means, for example, "motion, texture, visual appearance, and inner feelings and sensations" (Dingemanse, 2018: 3), as demonstrated in Table 1. Similarities across sensory modalities display a form of cross-modal iconicity (Ahlner & Zlatev, 2010; Jääskeläinen, 2013). *Interjections* – type of "response cries" (Dingemanse, 2018: 3) or exclamations – such as, *ooh, ouch, oops,* or *phew* are examples of words that are (indexically) associated with inner sensations and feelings.

The term *ideophone* is widely used as an umbrella term for "marked words that depict sensory imagery" (Dingemanse, 2011: 25). Ideophones are marked in the sense that they stand out "in terms of prosody, phonotactics and morphosyntax" (Dingemanse, 2018: 3), and reduplication is common – often occurring with phonemic variation (e.g. *ding dong* and *snip snap*) (Jääskeläinen,

2013). Finnish ideophones – also known as *expressive words* in the Fennistic tradition – for instance stand out by their atypical phonemic distribution and consonant clusters (ibid). As an example, word initial and word final consonant clusters are possible in Finnish ideophones, like in ***pr**uiskauttaa 'squirt'*, *tö**mps** 'thud'*, and some vowels are more common in these expressions than others (ibid: 13).

Table 1. Ideophones and phonestemes. Examples of English and Finnish ideophones divided further in types in regards to modality they describe, and examples of English and Finnish phonestemes.

Ideophones			Phonestemes	
Unimodal		Crossmodal		
Sound to sound resemblance (onomatopoeia)		Words that phonologically "depict" something in some other sensory modality than audition – often combining different modalities.		
English	- *buzz* - *snap* - *bang* - *pop* - *knock* - *meow*	English	- *dizzy* - *splash* - *sizzle, jiggle, wiggle*	English	Word initial *gl* = *shininess* - *glow, glisten* Word final */-ɘl/* = *small, fast, movement* - *jiggle, giggle, sparkle, drizzle*
Finnish	- *surina* - *napsahdus* - *pamahdus* - *poksahdus* - *kopsahdus* - *miau*	Finnish	- *humpsahdus* 'falling softly' - *lumpsahdus* 'falling in water'/'something slippery falling' - *molskahdus 'splash'* - *viuhahdus* 'swoosh'	Finnish	*hu = surprise, speed, airflow* *-humps, humahtaa*

Sound symbolism has increasingly received attention since the early 20th century, and experiments since then have consistently shown that some sounds tend to be associated with certain physical features, such as, shape, movement, size, texture, and colour (Ahlner & Zlatev, 2010). This kind of iconicity is often connected to the (pseudo) word pair *bouba* and *kiki* coined by Ramachandran and Hubbard (2001) in whose research the former was connected with roundness and waviness, and the latter with sharpness and pointiness by 95% of the participants. Similarly, in Finnish language research the phoneme /*i*/ has commonly been associated with tenseness, speed, and slipperiness, the /*o*/ and /*u*/ with slowness, softness, laziness and clumsiness, and /*œ*/ and /æ/ with unpleasantness (Leskinen, 2001 & Rytkönen, 1940a: 22 in Jääskeläinen, 2013). High front vowels tend to represent smallness, and deep back vowels and long vowels (i.e. vowel expansions) bigness (Jääskeläinen, 2013).

A sound symbolic phonemic unit that carries conventional associations to certain meanings is called a *phonestheme*. For example, in English *gl* is connected with shininess (e.g. *glow, glimmer,* and *glitter*), and in Finnish *hu* associates with surprise, speed and air flow (*hujahtaa – 'flop'*,

hulmuta – 'flow', hups – 'oops') (Jääskeläinen, 2013: 18) (see Table 1). Related to this, some sounds or segments express sensory meanings, for example, plosives as word final phonemes (plosive or plosive + *s*) in Finnish are used to imitate sudden sounds, for example, in words like *pläts – 'splash', kops – 'knock', tik – 'tick',* and *läps – 'slap'*, and word initial *l* in onomatopoetic expressions is linked to words that depict sounds connected with liquid (e.g. *läiskis – 'splash', loiskis – 'plash', lorina – 'gurgle'*, *lirinä – 'ripple',* and *litinä – 'squish'*) (ibid).

Although sound symbolism may exist on the level of single phonemes associating some sounds or sound combinations with certain characteristics that may affect the word semantics (as demonstrated above), they can of course be arranged in various combinations with other sounds thus forming into meaningful words (signs) (Hockett, 1960). This feature that allows the creation of virtually infinite number of different sound combinations from a limited set of distinct sounds at our disposal is unique for the semiotic system of language, and is referred to as *duality of patterning* (ibid) or *double articulation* (Martinet, 1984)[8]. Instead of "one sound one meaning", such duality affords the richness of the world's vocabularies and creation of evermore new signs for new phenomena (Trask, 2004).

Another and one of the most important features of the semiotic system of language is that it has a grammatical structure; all natural languages without exceptions have grammars (Hockett, 1960; Coseriu, 1985, Trask, 2004) allowing sentences to represent an open number of different events, with specific temporal and spatial features. Further, in order to create texts that form unified wholes rather than random utterances without relation to one another one needs devices that can be used to link clauses together to form *cohesion.* This kind of connectivity can be achieved, for example, by means of "syntactic conjunction and subordination (subordinating conjunctions, relative clauses), nonfinite verb forms, nominalizations, topic ellipsis" (Berman & Slobin, 1994: 19). These different structural devises are important for mapping causal and other relations between events in a discourse (cohesive devices will be explained in more detail in 2.4.3).

---

[8] Although many linguists treat the two notions more or less identical, it should be pointed out that Martinet (1984: 34) views *duality of patterning* and *double articulation* as two slightly different phenomena: "The point of view is not the same and the implications probably differ". He explains that "[d]uality of patterning just points to the existence of two different procedures, one for isolating phonemes and another for isolating significant segments in the utterance. Double articulation posits the existence of a process whereby a person who wants to communicate an experience unconsciously lets that experience be analyzed into a succession of elements, each one corresponding to one of the monemes [morphemes] of the language he will be using for his communication, and ordered according to the syntactic patterns of that language. Such a process results in what is called the first articulation. The second articulation, that of the vocal, perceptible face of the monemes into phonemes, does not result from a process performed by the man who is about to speak: as soon as the suitable moneme has been chosen, its vocal form, ready made, will be at his disposal" (ibid: 34-35).

2.3.2. GESTURES

Over the past decades, gesture research has gained much interest. Gestures can be described as "*expressive movements performed by the hands, the head, or any other part of the body, and perceived visually*" (Zlatev, 2015b: 458). Research has led to general acceptance that gestures exist in all cultures and that there are similarities in the way they develop in children across cultures. Much research in gesture studies focuses on so-called co-speech gestures, that is, spontaneous movements of hands, and occasionally head and torso, that are performed while speaking (Goodwin, 2003; Green, 2014).

The notion can be conceived either more broadly corresponding to what in some other fields may sometimes be called "body language", containing features like posture, facial expressions, different kinds of bodily movements, and gaze patterns, as well as pointing, etc.; or more narrowly as defined above, and typically used in connection with speech (Andrén, 2010). Gesture researchers usually consider gestures in terms of the narrower description of the term, and this definition is also followed in this thesis. To separate actions that are considered filling the requirements of gestures in the narrow sense from any other bodily actions Andrén (2010) draws a distinction between what he calls the lower limit and the upper limit of gesture (see Figure 5). This division separates volitional interactional bodily expressions considered as "gesture proper" – the semiotic system of gestures in our terms – from (a) those bodily movements or alterations of appearance that somehow either lack "some required degree of *volition* (such as blushing) or some required degree or kind of *semiotic complexity* (such as taking an object offered by another person)", and (b) signed languages (ibid: 13-14). That is, instrumental action and simple body language is placed below the lower limit for the reason that they are not composed of signs (see 2.2). Signed language, on the other hand, is placed above the upper limit, simply because it is one specific variety of the semiotic system of language, with large vocabularies, and complex grammars (Trask, 2004).

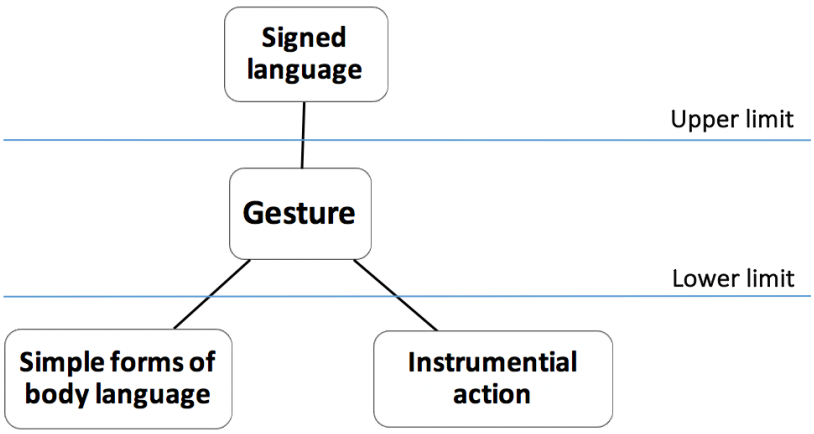

Figure 5. The upper and lower limits of gesture (adapted from Andrén, 2010: 13).

In the literature gestural action has often been analysed in terms of *gesture phrases,* which can further be divided into smaller segments: *preparation, stroke, (post-stroke hold)*, and *recovery* (Kendon, 2004). The stroke is what carries the gestural meaning in the gesture phrase, and since there can only be one stroke in a gesture phrase, identifying phrases can be done based on strokes on occasions when gesture phrases follow each other as a constant flow without returning back to the resting position (ibid).

Similar to language, there are many functions that gestures serve in an utterance, such as regulating the interaction, signalling intention, expressing representational content or structuring the discourse (Kendon, 2004). There is no one unified classification scheme of gestures due to their *multifunctional* character. However, many gesture scholars make a threefold division between a) *iconic* b) *deictic,* and c) *emblematic* gestures (ibid). This division reflects the traditional semiotic triad of a) similarity-based sign (icon), b) contiguity-based sign (index), and c) convention-based sign (symbol), as described in section 2.2. What is left outside this triad, however, is the type that functions to "mark out, punctuate or some other way make reference to aspects of the structure of the discourse, either in respect to its phrasal organization or in respect to its logical structure" corresponding to d) *pragmatic* function of a gesture sign (Kendon, 2004: 103). Thus, *pragmatic gestures* function on a different level than the other three types, whose function is referential to the content of the utterance rather than relational in regards to speech or the discourse. That is, pragmatic gestures do not refer to the content of the narrative, but rather they are, for instance, used to aid the interpretation of the utterance, structure it or regulate the interaction.

That gestures are multifunctional in character means that a single gesture can carry manifold functions at once, for instance, carrying pragmatic functions while being referential, depending also of its size and locus in the gesture space, its dynamics, and temporality in relation with speech (Payrató and Teßendorf, 2014). Following from such characteristics, a clear-cut universal gesture typology is difficult to draw, as such would not serve to take into account all the necessary aspects of the phenomena in each situation. Therefore, researchers have typically focused on specific attributes of gestures that they have taken under analysis creating a typology for that purpose and leaving out secondary aspects that are not predominant in the context (Zlatev, 2015b). Following the threefold division of the semiotic triad of icon, index, and symbol, the following gesture categories are introduced next in more detail: *iconic gestures, deictic gestures,* and *emblematic gestures.* The category of pragmatic gestures is then added to this triad to take into account those gestures that do not fit into this division.

Iconic gestures are a class of gestures "where there is resemblance between the movements of the whole body, or parts of it, and properties of intended actions, objects or whole events"

(Zlatev, 2015b: 461)[9]. This class can further be divided into *enactments* and "*symbolic" gestures*, where in the former the body of the gesturer matches the action that is represented, and in the latter it does not, like when moving fingers do not represent fingers that move, but are taken to represent, for example, moving legs (ibid). Arguably, the former has a higher degree of primary iconicity, as shown in Figure 6. The "symbolic" *non-enacting* gestures may vary in specificity, so that hands can be used to represent something specific and concrete like an animal (or moving legs), or something non-specific, an abstract entity such as art (ibid).

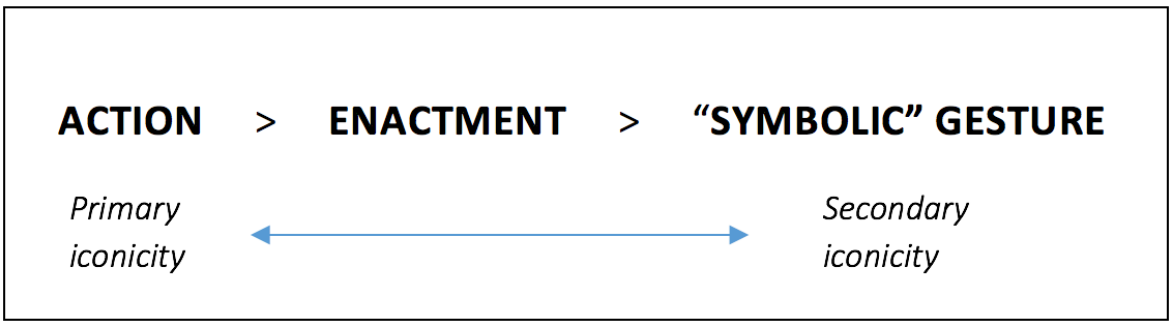

Figure 6. The level of iconicity from enactment to symbolic gestures.

McNeill (1992) assigns iconic gestures a *viewpoint* distinction that the speakers/gesturers use to produce the movements: *character viewpoint (CVPT)* or *observer viewpoint (OVPT)*. In the former the speaker enacts the protagonist as though being "inside" the character; whereas the latter viewpoint takes the observer perspective, so that, for example "two hands can be used to mold the shape of a protagonist or a finger can be used to draw a path that a protagonist is moving along" (Debreslioska et al., 2013: 436). A mixture of the two viewpoints is also possible, for example, "when the hand is representing the trajectory of an inanimate entity, such as a ball that is flying toward a person, whereas the body of the speaker represents the person who is being approached by the ball" (ibid: 450). While not fully identical, this division corresponds to that between enactments and "symbolic" non-enacting gestures, described above.

Different techniques with gestural articulators can be used to display the abovementioned gestural perspectives in iconic gestures. Müller (2014: 1690) divides these different techniques into the categories of *acting, molding, drawing*, and *representing*. She proposes that these techniques "address iconic motivations of gestures" but are not necessarily only linked to iconic gestures per se (ibid: 1691). As the names suggest, in the acting technique hands or the body is used to re-enact the

[9] This definition of iconic gestures differs from that given in McNeill (1992: 109), where they are described as "imagistic" and "unwitting".

depicted movement, and in the molding technique the hands are used to "mold" an object, for example, a ball by forming a round shaped item between two hands as though holding the ball with this wrapping grip. In the drawing technique one or both hands, or fingers are used to "draw" the outline of a shape in the air or against background, or to display a trajectory, and in representing technique a hand represents an object, for example, when hand with fingers spread out wide stands for an antler (ibid). The combination of "acting" and "molding" techniques is similar to the enacting category of Zlatev (2015b) and the character viewpoint (CVPT) of McNeill (1992), thus involving more primary iconicity, while the other two are similar to the "symbolic" non-enacting category of Zlatev and observer viewpoint (OVPT) of McNeill, thus involving more secondary iconicity.

Deictic gestures are those that "indicate or individuate an external target for an addressee, and include not only different types of pointing, but also acts which bring an object to the attention of the addressee (showing, giving, requesting)" (Zlatev, 2015b: 461). There are also abstract uses of deixis. Abstract pointing includes the use of gesture to refer to an imaginary object instead of something in the vicinity (Andrén, 2010).

Emblematic gestures, or emblems, are the most language-like gestures. They are conventional in both form and meaning, and thus they clearly qualify as symbols (Efron, 1972; Andrén, 2010), as defined in Section 2.2. Emblematic gestures can be described as "those nonverbal acts which have a direct verbal translation, or dictionary definition, usually consisting of a word or two, or perhaps a phrase" (Ekman & Friesen, 1969: 63, quoted by Andrén, 2010; 46). Just like linguistic signs, emblems differ across cultures in number, and form-meaning associations (Kita, 2009). Examples of emblems shared in many Western cultures are, for example, the PEACE sign (formed with index and middle finger), the OK sign (circle formed with thumb and index finger), or GOODBYE wave.

When discussing pragmatic gestures, the function these gestures serve in an utterance is usually highlighted. The pragmatic function is itself multifunctional, and depending on the situation a gesture can, for example, have a) *performative* function – indicating whether the speech act is to be understood as a request, offer, rejection, or something else, b) *modal* function – indicating how what is said ought to be interpreted (e.g. expressing negation), c) *parsing* function – punctuating or structuring the speech (Kendon, 2004). Type of parsing gestures referred to as *beats* may be used rhythmically to help organising crucial plot information, thus emphasising the key aspects of the narrative (McNeill, 1992).

2.3.3. Pictures

Pictures are typically visually perceived iconic signs (Sonesson, 2011). According to Sonesson

(2011: 173) "[a]n icon must be understood as iconicity […] plus the sign function". That is, if there are two things that resemble one another, one (expression) must be understood as standing for the other (content) and not only have a meaning on its own right. The picture is an example of a primary iconic sign (Sonesson, 2010).

A second feature of the pictorial system is what has been termed *resemantisation* (Sonesson, 2011: 173). Lines and dots that a picture is composed of are meaningless until they find their place in a picture that they are part of. That is, once put together the parts of the picture take up aspects of that overall meaning, such as a hand and leg of a character. This is distinct from duality of patterning (see Section 2.3.1.) as individual phonemes do not take up aspects of the whole. For example, the English word *cat* is formed by combining the phonemes /*k*/, /æ/, and /*t*/, but the single phonemes will still be deprived of any feline attributes.

Further, Sonesson (2011) analyses the picture sign in terms of Husserl's notion of *pictorial consciousness*, which he develops further involving the following four levels: (a) *picture thing* (the physical picture), (b) *picture object* (what is seen "in" the picture thing), (c) *picture subject* (what is seen, but now with its "proper"[10] colours or details) and (d) *picture referent* (the thing in the world that is photographed or drawn) (ibid: 174-175, emphasis added) – for demonstration see Figure 7. Thereby, any critique, for example, addressed to the poor condition of the picture would be a reference to the material quality of the picture thing, but if the critique is instead addressed to the failure of the picture to present resemblance of the person or thing it portrays (i.e. the picture subject), the issue would be with the picture object. The picture referent, then, can always be found in the perceptual world if the referent indeed exists, as is the situation with land marks (e.g. buildings) and living creatures photographed or by other means portrayed in a picture thing. Storyworld characters, however, have a different ontological status compared to those in the perceptual world. Such is the situation with the story this thesis is to present in the forthcoming chapters – a story, whose main characters are *a boy*, *a dog* and *a frog*, who are examples of picture objects. Even though they do not have referents (being imaginary) they do, however, have picture objects and subjects, for they are characters in a fictive story.

[10] Even photographs that virtually have "an expression plane which is tautologously related to its content plane" are never fully able to reproduce reality in its true colours and light conditions, especially with the technology that enables enhancing the desired details and fade the undesired (Sonesson, 2015: 432).

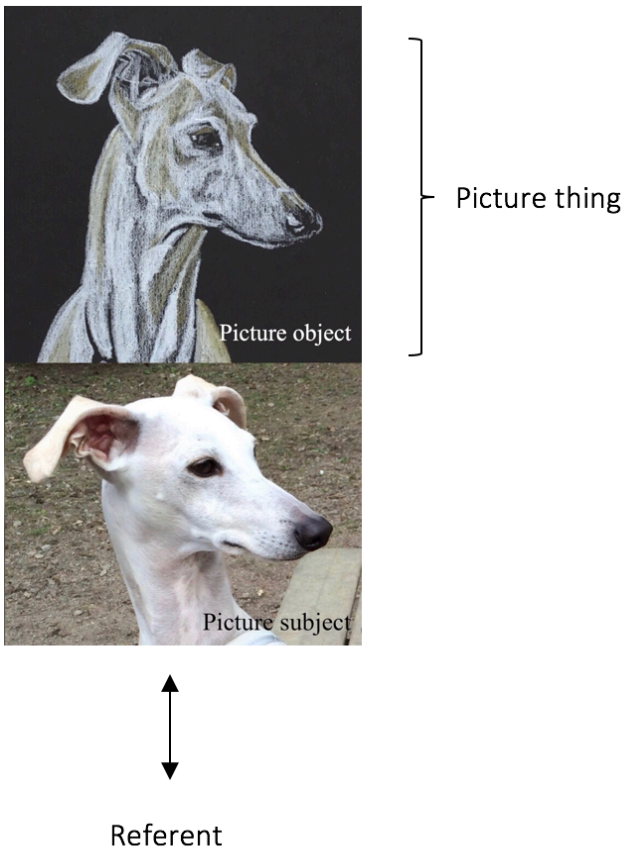

Figure 7. Pictorial consciousness as manifested in relation of *picture thing* (the art work), *picture object* (what is displayed in the drawing – a dog), *picture subject* (the dog in his real appearance), and the referent (Italian greyhound named Chili, who exists in the perceptual world) (conforming the model of Sonesson, 2010: 48).

As explained in Section 2.2, any concrete sign can be based on a combination of different grounds (i.e. iconicity, indexicality, and conventionality), but what determines its type is the predominant one. Therefore, although a picture is clearly an instance of an iconic sign, it can also have grounds of indexicality and conventionality[11]. Pictorial semiotics has even sometimes gone as far as postulating indexicality as the predominant character of photographs: "a trace left behind be the referent itself" (Sonesson, 2015: 441). However, since photographs and other pictures are above all iconical (similarity as the ground) it is only through iconicity that they can manifest their indexicality (ibid). The indexical relation between "the photographic expression plane", and its object(s) in the perceptual world at the moment the image was created is in this situation that of contiguity (ibid: 450). Furthermore, in a series of pictures time contiguity can be used to express indexicality by means of a degree of change in otherwise identical pictures (or otherwise connected by theme, as explained below) – only requiring that a comparison is possible to make. A change

[11] Conventionality clearly manifests itself in different pictorial genres, but also in the equipment and styles used in the production of those images – nuances and contrasts (Sonesson, 2015).

such as that in brightness, according to Sonesson (2015: 432), is based on "biologically grounded" reaction to phenomena where a change indicates progression of time (e.g. break of dawn) or arrival or passing of a storm. Just as the change in light or colour in pictures may be an index of time contiguity, so can seasonal change in scenic pictures, or even in family photographs (e.g. changing décor of domestic environments). A change detected in picture objects can thus also be an index of maturing and aging of referents and time contiguity may furthermore be reflected through fashion and design. Indexicality may also be based on contiguity in terms of causality: a picture of meal ingredients followed by a picture of a cooked dinner, or an image of a child kicking a ball and another of a broken window; or on factorality (i.e. part-whole relation): a picture of a pregnant woman and an ultrasound image of the fetus in the mother's womb.

2.3.4 Polysemiotic communication

Language, gestures and pictures are three central semiotic systems for meaning making at our disposal, and they may be incorporated in parallel for the creation of the desired effect – with the different expressive potentials of each system related to the semiotic properties discussed in Section 2.2 – with the particular affordances of subsystems like speech or writing. Thereby, following Green's (2014: 10) proposal, multimodality in this thesis is used to refer to "communicative ensembles or utterances that draw on more than one of the main sensory modalities" – for example, vision and hearing, as in typical face-to-face interaction that combines speech (which itself can also not only be heard but seen) and gestures. This approach is in contrast to Kress' concept of multimodality based on rather vaguely defined "semiotic modes" (Kress, 2010: 166), which are difficult to compare (e.g. colour, gaze, speech, posture etc.) (see footnote 2).

The significance of this kind of polysemiotic and multimodal orchestration becomes evident in face-to-face interaction, but also elsewhere in our communicative environment. For example, it can be seen in the design of textbooks, websites, and road signs, which may have pictures (or video/moving image) and sound in addition to written language. Until recently, the focus of communication research has predominantly been system specific, traditionally concentrating on written or spoken language. However, the importance of research that takes into account the interplay of different semiotic systems utilising one or more sensory modalities has increasingly been acknowledged, because focusing on one system in isolation is not sufficient in order to understand the richness of human communication holistically (Green, 2014).

As is evident from the previous discussion, the semiotic system of gestures is often tightly linked to the system of language when manifested as speech. In the present study gestures and speech are understood to be closely interlinked but distinct semiotic systems instead of regarding the

two together constituting one "single system" (e.g. McNeill, 1992) – following the lead of Kita and Özyürek (2003), Kendon (2004), and Zlatev (2015b). In this synergetic relationship gestures and speech interact closely with each other to create a richer form of *utterance* through a combination of two qualitatively different representations: spatio-motoric, and linguistic (Kita & Özyürek, 2003). Yet, despite this tight interaction, the organisation of the two modes of expression is flexible in a way that co-speech gestures can be "orchestrated" differently depending on circumstances (Kendon, 2004). Therefore, gestures are not to be regarded as "automatic" and "unwitting" like has sometimes been argued in the gesture literature (e.g. McNeill 1994), but as manifesting "deliberate expressiveness" (Kendon, 2004: 15), even if the speaker/gesturer is not fully conscious of this deliberateness.

The two semiotic systems are used in parallel so that not only are gestures and speech semantically and temporally co-expressive in terms of content, but gestures can also reflect discursive (interactional) structures (Kendon, 2004; Gullberg, 2016). Gestures may facilitate the representation of location and space while speech might be more effective in representing categorical relations among entities (Wagner Cook, 2014). Nonetheless, in cases of disfluency in lexical retrieval, when speech stops, usually so does gesturing (Gullberg, 2006; Graziano & Gullberg, 2018). Rimé and Schiaratura (1991) point out that gestures tend to be concentrated around the first word of a fluid utterance and around the first word after hesitation. However, the gesture unit is known to start slightly ahead of the lexical item that it marks (Kendon, 2004), which could support the idea of facilitative role of gestures on speech production. In effect, co-speech gestures may reflect strategies in structuring and planning speech, and thus decrease cognitive load by shifting it from verbal working memory to other cognitive systems or external representations (Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001; Wagner, Nusbaum & Goldin-Meadow, 2004). Different theories concerning the role of gestures in speech thus highlight, once again (see Section 2.3.2.) the multifunctional character that gestures have in human communication.

What could explain the close synchronisation of the two semiotic systems in human communication according to Zlatev (2007: 263) is the concept of *bodily mimesis* – the use of the body for creating and communicating representations:

> This [cognition based on mimesis] allowed our ancestors to use their bodies to perform elaborated actions that others are observed to be doing (imitation), to represent external events for the purpose of communication or thought (pantomime, gesture) and to rehearse a given skill by matching performance to an imagined goal.

According to this theory, a new form of bodily cognition in early hominids – based on mimesis – preceded language, for even though mimesis has the properties of representationality and

accessibility to consciousness, it lacks the following two central features of language: full conventionality and systematicity. This would indicate that *imitation*[12], *pantomime*, and gesture are not dependent on language, but rather, language was built atop of them (ibid). With empirical support Zlatev (2007: 245) introduces the notion of *mimetic schema* (not to be mistaken with *image schema*[13]), described as "body-based, pre-linguistic, consciously accessible representations that serve as the child's first concepts". Mimetic schemas can be explained as representations of real or imagined everyday bodily actions and events, which all have their specific "feel" as a result of being experiential, making them "an aspect of phenomenological embodiment" (ibid: 268). That is, mimetic schemas are specific; each schema is a generalisation of a certain bodily act, such as, KISS, RUN, EAT, KICK, which act as "grounds", motivations for linguistic symbols.

According to Zlatev (2014), iconic gestures emerge in ontogeny as overt mimetic schemas. Since they arise through imitation and social interaction, mimetic schemas are also largely culturally shared, typified actions rather than individual, private representations. Consistently with this, Kita (2009) argues that gestures reflect the way we conceptualise events in communication in a given culture. That is, gestural presentations also depend upon the syntactic and lexical resources available in the given language, which, in part reflects the corresponding culture.

2.4. Stories, narratives, and narrativity

The field, whose main objective is to study narrative structures is called *narratology*. However, narratives are a topic of study within a variety of disciplines from film studies to pedagogy, and are thus a relevant topic for cognitive semiotics. Not only are narratives a way to organise spatiotemporal events in a chain of causal episodes having a beginning, middle, and an ending, but they also contain evaluations of the nature of these events and indicate the reasoning behind these evaluations (Kajannes & Kirstinä, 2000). Further, narratives can offer fruitful material about sociocultural aspects in language use (Labov, 1972), language and gestural acquisition, cognitive development (Berman & Slobin, 1994; Gullberg, 2006), identity construction (Gal, 1978), and face-to-face interaction (Kallio, 2010). This section starts by introducing the main concepts, and then

[12] "Imitation (or 'copying') refers to the skill of appropriating actions performed by others into one's own repertoire." (Andrén, 2010: 51)

[13] "The concept image schema was initially defined as 'a recurring dynamic pattern of our perceptual interactions and motor programs that gives rise to coherence and structure to our experience'" (Johnson 1987: xiv as cited in Zlatev, 2014: 3). The term appears to be rather polysemous often used in Cognitive Linguistics to describe rather abstract structures – like PATH – that have traditionally been thought to "provide a *ground* for linguistic meaning" (Zlatev, 2005: 3, emphasis added). Whereas mimetic schemas are typified actions, image schemas are hypothetical constructs (Zlatev, 2014).

moves on to describing and comparing different narrative modes relevant for the current thesis, and the general narrative structures and devices used in narrative construction.

2.4.1. Concepts

The ways the concept of narrative is defined in the literature are many, but a minimal definition is that of *a verbal or non-verbal* (e.g. pictorial or gestural) *display of two or more events* (Prince, 2008: 19). Such events are usually in chronological order, and causally contribute to later events, and can be represented using different semiotic systems (Sanford & Emmott, 2012). Thus, when used for producing narratives these semiotic systems function as *narrative modes*. A conceptual distinction between *story* and *narrative* is important to (re)make at this point, as even though they have been treated as synonyms in some of the literature, a distinction has been made in others. Berman and Slobin (1994) and Clark (2004) define a "story" as the product, and "narrative" as the process of telling the story. Elsewhere, in classical narratology, these (though perhaps not strictly with one to one correspondence) have also been termed as *fabula* – story in its chronological order of events – and *sujet* – narration with its narrative strategies employed that may include flashbacks or "flashes" forward (Bacon, 2000).

In Section 2.3.1 Coseriu's (1985) perspectival distinction between language as activity (*energeia*), knowledge (*dynamis*), and product (*ergon*) was introduced in regard to language, but it can also help resolve the ambiguity of the concepts here. *Narrativity,* could be seen corresponding to knowledge – connecting narratives (the free activity of producing or understanding a story) to stories (i.e. narrations in particular media). This is consistent with proposals to consider narrativity in terms of structures of consciousness, even preceding any expression (e.g. Ranta, 2011). Menary (2008) uses the term *embodied narrative* to refer to narratives that reflect the experiences of the "lived body" in the phenomenological sense, but this corresponds in our terminology to narrativity. Thus, I use the term here to refer to retold narratives in which the body participates as a semiotic resource in the narrative (i.e. in the *telling* of the story).

2.4.2. Narrativity and narrative structure

As mentioned in chapter 1, it has been argued that language may have initially evolved as a means for expressing narratives. As proposed by Donald (1991: 745), "day-to-day storytelling in a shared oral culture eventually produces collective, standardized narrative versions of reality, particularly of past events; and these become what we call the dominant 'myths' of a society".

On the other hand, Ranta (2011) argues that narrativity is part of our embodied cognition, even without language. According to him the ability to think in terms of stories is essential for us to make sense of the world around us, to manipulate it, to foresee changes, and understand causal

relationships. Basing his view on notions from cognitive psychology, Ranta (2011: para. 3) proposes that cognition involves storage and retrieval of action schemas (he calls these simply *narrative structures*) – that "incorporate generalized knowledge about *event sequences*, such as the order in which specific events will take place; causal, enabling, or conventionalized relations between these events, and what kind of events occur in certain action sequences." These narrative structures are acquired through earlier experiences and thus are largely culturally shared, with some idiosyncratic variation. They need not be first-hand experiences, but can also be acquired indirectly through narratives in one or another semiotic system. Stereotypical examples of these schemas could be, for example, the event of going to a funeral and wearing black, or buying groceries and paying with a credit card at the counter. Terms commonly used for this type of sequences of categorisable actions in cognitive psychology are for instance, *scripts, frames,* or *event schemas* (Mandler, 1984). All these can be understood as structures of narrativity, as defined in the previous subsection.

Similarly, Fludernik (1996: 12) views narrativity as a natural product in the "framework of *human embodiedness*" and thus considers it being tightly connected to experientiality. In this framework she distinguishes four levels that operate in narrativity. On the first level – referring to the *what* – the "reader" intuitively views narratives through her/his pre-existing narrative structures (such as those described in the previous paragraph) – "parameters of real-life experience" – that are key for understanding agency as "goal-oriented process or reaction to the unexpected, the configuration of experienced and evaluated occurrence, and the natural comprehension of observed event processes including their supposed cause-and-effect explanations" (ibid: 43).

In the second level – that refers to the *how* – the "real-world" schemas of *telling*, *viewing*, *experiencing*, and *acting* are aiding in the process of entering the story (ibid: 43). In addition to understanding goal-orientedness, this level initiates the "processuality of event and action series" (ibid: 44), as well as reflection, and the processes of introspection and memorising can be intertwined with the act of telling of the story. The personality of the receiver of the narrative – to whom the narrative is directed to – affects the narrative process by creating a communicative situation.

On the third level culture has the largest role, for this level entails knowledge of different story telling frames (narrative situations and the way they are structured); for example, narrators and their audience, and the interaction between them, but also knowledge of different narrative genres. These third level frames are nuanced extensions of those on the first level; for example, the knowledge about the dynamics of different conversational exchanges one can participate in on a bus or on a date.

All these three levels provide "cognitive tools for the interpretation" (ibid: 45) of narratives, but it is not until the final level that we reach different strategies for interpretation. This fourth level refers to narrativisation – how the receiver of the narrative uses different categories in the first three levels to *naturalise* and make interpretable what is unfamiliar about the narrative. The notion of naturalisation is not unlike *the theory of remembering* put forward by psychologist Frederic Bartlett in 1932, in which memory is considered a process of reconstruction, whereby the existing schemas have an effect on how more unfamiliar occurrences are remembered.

Another dimension of narrative structure concerns the various components or aspects of a story or narrative. According to analysis provided by Sanford and Emmott (2012) a narrative is typically built up of the following aspects (though, all of them need not be included): *setting*, *theme*, *plot*, and *resolution*. The *setting* can include, for instance, the characters of the storyworld, the location(s) where the story events happen, and also the passing or progression of time. The *theme* of a narrative expresses what the story essentially is about; it can be in the form of a goal to be achieved by the characters, the teaching of the story, or the issues story raises, for example, death, spiritual growth, or loneliness. The *plot* is an important aspect of the story, the red thread that is carried through the narrative. It is the sequence of events that leads up to the *resolution* – that is, the outcome. The plot structure is divided in three core components[14] in the study of Berman and Slobin (1994: 46): *onset*, *unfolding*, and *resolution.* That is, for a plot to be meaningful, it needs a beginning (the onset), an end (the resolution), and events in the middle that lead up to the outcome (the unfolding).

2.4.3. Narrative modes

Narrative research has predominantly focused on language (in literature and oral discourse), or more recently, film and television, where events are represented in a time sequence (Ranta, 2011). The importance of oral narratives is undeniably immense in human cultures, but other semiotic systems that are commonly used in modern societies to "tell" a story are those of gesture (e.g. in pantomime), and pictures, including picture sequences (Kress, 2010). This latter semiotic system, and its corresponding *narrative mode* is compared in the experiment described in this thesis together with semiotic system of language and its corresponding narrative mode, and thereby I will demonstrate here each of their expressive potentials and restrictions when presenting the same story content. These semiotic systems differ from one another in regards to how they can be used to do specific semiotic work, having such different potentials for meaning making, as discussed in

[14] Berman and Slobin (1994) count resolution as part of the plot structure, whereas Sanford and Emmott (2012) count the events that lead up to the resolution as plot.

Section 2.3.

The semiotic system of language is characterised by the sequential arrangement of elements in time, which provides the semiotic logic in this narrative mode (Kress, 2010). Verbal narratives (spoken, signed or written) are linear structures that require "a temporally successive perceptual process" (Ranta, 2011: np). The narrative property that facilitates this process is the plot, which may lack a chronological order, or strictly causal entailments. In language this unfolding of narrative information involves "linguistic *cohesion* on the micro level of individual clauses and adjacent clauses, and thematic *coherence* on the macro level of plot organisation" (Berman & Slobin, 1994: 40, emphasis added). It is crucial to know who does what to whom when and where, which can be achieved by means of reference tracking. Languages offer several different mechanisms to indicate the information status of referred entities depending on their accessibility (givenness) (Hickmann & Hendricks, 1999). In English, for instance, when a referent is introduced for the first time, a full lexical form with an indefinite article is used (e.g. *a boy*). Less full form of entities mark information that is already accessible to the listener (e.g. the pronouns *he*, *him*, or *his,* or the definite nominal *the boy* when the referent is reintroduced), and "zero anaphora", as in (5) can be used when the referent is presupposed, marking the strongest continuation or relationship to the topic (ibid).

(5) *He heard a voice and _ peeked over a log* in the pond.

Another way to achieve cohesion in language is with help of connective devices, which can be used to link discourse segments together. Halliday & Hasan (1976: 244-270) provide a scheme of conjunctive relations in five categories depending on their function in the clause:

- *additive*: reflecting additive relation (e.g. *and, or, nor,* or *above all*)
- *adversative:* reflecting opposing relation (e.g. *although, but, yet,* or *instead*)
- *causal:* reflecting causal relation (e.g. *so*, *because, for this reason,* or *consequently*)
- *temporal:* reflecting temporal/successive relation (e.g. *then, next, meanwhile,* or *first)*
- *continuatives:* a rather diverse category including, for example, *well, anyway,* and *of course* which bring cohesion in communication in different ways by indicating a response to expressed events, as in (6), reflecting attitudes or rhetoric, or dismissal of the previous utterance, as in (7).

(6) *They went to see if it could be behind the rock.* ***Well*** *it wasn't there either.*

(7) *Perhaps they were sleeping.* ***Anyway*** *the window was open.*

Clark (2004: 462) argues further that "[n]arrators don't construct a narrative *simply* to fit a story, selecting their words, phrases, and rhetorical devices to express the elements of a pre-determined conceptualization of events." Rather, the content of a verbal narrative is limited to the

specific aspects that the narrator *chooses* to describe. Consistently with this, Sonesson argues that (2014: 274)

> The imperfection of verbal language, which is at the same time its great strength, is that it cannot render anything, temporal or spatial continua, or whatever, as anything but sensate qualities[15] – that is, in the form of a limited set of properties abstracted from the wholes of which the world of our experience is made up.

Thus, while bringing some pieces of information to the attention of the listener, any verbal narrative unavoidably leaves something out.

There are several devices that can be used in speech in order to bring desired aspects of the story in focus while leaving less important aspects in the background. Paralinguistic devices such as prosody are one example of an effective foregrounding strategy (i.e. means to bring something in focus) in spoken narratives (Kress, 2010). Sound is a rich element that offers a possibility for a considerable variation in energy, which can be used to stress elements of choice, accentuate words, or produce rhythm to organise speech. Variation in pitch produces intonation that is used, for example, to indicate attitudes or mark different kinds of phrases (i.e. questions or statements etc.), but also to *frame* – that is, mark if produced information units are new or given, and to tie these together into larger level coherent units, a frame (ibid). Framing is essential for meaning making in narratives, and all narrative modes have their own unique ways of doing so (ibid). Other means that can be used for foregrounding (and backgrounding) elements in a spoken narrative are repetition, speed of narration, and the order in which information is presented, but also, by explicitly referring to events that are not happening, and the qualities that are not present in the storyworld may emphasise some aspects in the story that are (Prince, 2008). In storytelling the use of formulaic expressions such as, *once upon a time there was* and *they lived happily ever after* is also a way of framing, to mark the beginning and ending of a story, and to indicate the type of narrative genre (Ortactepe, 2012).

From a cognitive perspective, Fludernik (1996) argues that what is significant in any narrative is the presence of a conscious actor and experiencer on some level of narration. This factor brings us back to the notion of (anthropomorphic) experientiality, which is an essential part of all narratives from oral everyday storytelling to literary prose. Sometimes the experience reflects the consciousness of the protagonist – "the reflector-mode" (experiencing) – but often in the focus

[15] Sonesson (2014: 275) argues that sensate qualities "result from the extraction of some – but ideally the most important – properties of spatial continua. […] I would contend that pictures are limited in the number of sensate qualities they can convey, but not as limited as language."

is the consciousness of the narrator (narrating) (ibid). Different narrative genres have their structural and stylistic conventions, and Fludernik (1996) lists different types of oral narratives according to their level of experientiality. Of those enlisted I will briefly introduce two types that stand towards the opposite poles of the continuum of experientiality: *narrative report* and *narratives of personal/vicarious experience*. Report merely is a summary of events and action sequences rather than a narrative that develops into an actual story – that is, its main function is simply to provide information. This type of narrative is built upon "second-hand experience or on a summary of first-hand experience rendered non-experientially" (ibid: 71). Reports tend to be objective descriptions of action and event sequences followed by the resolution, effectively offering "the 'point' of the story" (ibid:71). Consequently, Fludernik (1996) links the reliance in connective device *then* with narrative report style. Unlike reports that correlate with distance, experiential storytelling (i.e. narratives of personal or vicarious experience) is associated with "experiential value" either from first person or third person perspective (through empathy or "*perceptual* focalization")[16], yet without adopting "'omniscient' stance" (ibid: 75-76, original emphasis). This type of narratives are "structured by means of an interaction between a narrative frame, a plotline level of the proper story, and an off-plotline level of embedded orientation and narrational evaluative or explicatory commentary" (ibid: 80). So, experiential storytelling becomes discourse-like story depiction with reflective elements. Fludernik (1996) links this narrative type with the use of connective device such as *so* that causally and thematically links together story events.

As pointed out several times, pictorial narratives have traditionally received much less attention in narrative research than verbal narratives (Sonesson, 1997; Ranta, 2011). Information in the semiotic system of pictures is typically received via sight organs, and all of it (in still pictures) is available simultaneously, unlike in the system of language. This characteristic allows a much quicker perception compared to language, which is why (highly conventional) pictures are frequently used for example in traffic signs (Kress, 2010). Since there are no words in pictures to describe events and name characters, other means are used to present events. The semiotic system of pictures, as pointed out in Section 2.3.3, is predominantly grounded in iconicity rather than conventionality. While language uses the logic of time as a resource, still pictures use that of space. As a narrative mode pictures show elements that may not be present in a narrative presented by linguistic means, such as, the location or size of objects, and the overall space. Thus, framed space is one of the affordances of pictures, as all elements are present simultaneously and arranged in the space in a way that conveys meaning about the relations between entities. Consequently, *transitivity*

[16] *Focalisation* here meaning the access to the consciousness of the character (Fludernik, 1996).

– that is, the relationship between agent and patient (i.e. who does what to whom) – is often directly given or interpretable in pictures (Sonesson, 1997). Furthermore, while in language narrative is necessarily carried linearly, pictures on the other hand afford much more freedom in the sense how they can be "read", for example, they can be examined horizontally from left to right or right to left, diagonally or top down (Seppä, 2012). *Colour* is a powerful resource for framing in pictures, and can be used to highlight specific elements in pictorial narratives in a similar manner than some linguistic techniques can be used to highlight elements in speech (Kress, 2010). Colour can thus be used to foreground elements in a narrative – bringing elements in focus.

The ability of static pictures to represent whole events and "tell stories" can be understood on the basis of the narrative structures explained in Section 2.4.2. According to Ranta (2011), a single picture has the ability to activate action schemas in the viewer that have been acquired through earlier life experience. He argues that "[p]ictorial material is frequently and intentionally produced in order to trigger stories or at least to give rise to narrative hypothesizing" (ibid: last para). In other words, how a static picture may invoke an action or event schema is explained by the selection of the most arrested moment in the desired action sequence, which may trigger activation of a bigger mentally stored event schema (ibid). Furthermore, details and peculiarities in expression may evoke viewers to imagine theories around the triggered action schemas. As an example Ranta (2011) takes Edward Hopper's painting *Automat* (1927), which depicts a well and warmly dressed woman drinking coffee in a restaurant, which instantly evokes a going-to-a-restaurant schema. Painting's details, however, infuse a series of questions on an attentive viewer: Why is she only wearing one glove? Is she going to meet someone special or is she coming from somewhere? Is there sadness in her life? Is the depicted moment late at night or on a gloomy winter or autumn evening? Such protentions (what could happen) and retentions (what just happened) that Sonesson (1997) connects to the notion of time consciousness (Husserl, 1936), discussed in Section 2.2, is a characteristic property of most pictures.

Ranta (2013: 8) briefly distinguishes two other types of pictorial narratives that are different in regards to narrative and temporal aspects from those of static pictures: *continuous narratives*, and *narrative series*. Continuous narratives were represented already in historical works of art, where a single painting may show several different persons and events in the same pictorial space. In this type of narratives different phases of an event series are represented simultaneously in a single canvas. Temporal continuum is more directly represented in such a pictorial narrative, and the idea of instantaneous nature of static pictures is less applicable with temporal processes being involved in the perception of multiple events. Narrative series, then, are most clearly represented in strip cartoons, where static pictures depicting different scenes or events are linked to form a narrative

(ibid). These distinct pictures have a fixed reading order – often horizontal or vertical – and temporal continuum (indexical relationship) thus is more directly represented and linear (Cohn, 2013).

The picture book *Frog, Where Are You?* by Mercer Mayer (1969) – famous in narrative research – could be seen as a narrative series type of a pictorial narrative. It is not a strip cartoon, but a wordless picture book where each picture represents an event and together they form a narrative that has a beginning and an end that are connected by events in between. When skimming the book, the first picture lets the viewer imagine a storyworld set in a bedroom in the evening with a boy, a dog, and a frog that is kept in a jar. When flipping the page to the second picture, the viewer does not create a new storyworld, but rather, adds a next scene to the previous one – a scene where the setting is otherwise the same but later at night (the boy and the dog are asleep). The successive pictures show changing settings of the *same* storyworld where the viewer is expected to add attributes and motivations that are not directly displayed in the pictures (Clark, 2004). The "frog story" has been widely used as a source of data to study narratives, especially from the perspectives of language acquisition and cross-linguistic comparison. Since the book does not contain language, the differences between produced verbal narratives based on the book can easily be attributed to specific features in the language, culture, age, and storytelling skills of the participants. The book is fruitful material also in gesture studies, because such series of drawn depictions, according to Clark (2004: 465), are an instance of "mimesis" – understood in the general sense of iconic representation – which we saw in Section 2.2 is also characteristic of the semiotic system of gestures.

2.4.4. Embodied narratives

As was noted in Section 2.3, what we make use of in our everyday lives for meaning making is the ensemble of semiotic systems as a whole. Face-to-face communication is polysemiotic when the different semiotic systems act together in a way that, for example, speech names and gesture points or enacts. Whereas language is, as we saw, based on logic of time, and pictures on that of space, in gestures both of these logics are combined. The logic of time in gestures can be seen in that "there is sequence in time through movement of arms and hands, of the head" (Kress, 2010: 81); and logic of space is presented in the movements located and taking form in speaker's gesture space. Thereby, gesture locates an entity in the gesture space indicating, pointing out, or displaying, while at the same time speech names, making it explicit.

Many studies have shown that ideophones also frequently occur in conjunction with iconic gestures (Kita, 1997; Clark, 2004; Dingemanse, 2018). As pointed out in Section 2.3.1 sound

symbolism is a reflection of unimodal or cross-modal iconicity in spoken signs, when they represent sounds or other sensory events or features of animate beings, things, and action. Ideophones have even been regarded corresponding to "depiction", for rather than giving a description of their content they let the listener to imagine the scene and in a way carrying "gestural properties" of their own, often further being combined with actual gestures (Dingemanse, 2018).

The interplay of the two narrative modes acting together is also evident when it comes to managing coherence relations in spoken narratives. It has been demonstrated that reference tracking is a bimodal and polysemiotic phenomenon (Levy & McNeill, 1992; Kendon, 2004; Gullberg, 2006). Gesture space can be used effectively in order to create cohesion in a narrative. It can be done by "continued or recurring gestural patterns – handedness, hand configuration, or specific spatial area – associated with consistent visuospatial imagery or referential content over a stretch of discourse" (Gullberg, 2006: 158). The iconic and indexical aspects of gestures are typically used to achieve this association, which is systematic and recurrent between speech and gestures (Levy & McNeill, 1992). In effect, when a referent is introduced, it is often anchored in a specific location in the gesture space with a deictic or iconic gesture. This referent-specific locus is typically maintained throughout the discourse of narrated events, and when more referents are introduced, they may each be given a locus of their own, which makes referent tracking easier for the interlocutor (Levy & McNeill, 1992; Gullberg, 2006; So, Kita, Goldin-Meadow, 2009). This method turns the gesture space into a type of visual "map of discourse" where the speaker juggles between these loci to maintain the established map of referents throughout the narrative the (Gullberg, 2006: 159). Here we can see the interplay of two semiotic systems where one complements the information provided by the other.

2.5. Translating between semiotic systems

Finally, we come to the term that features in the subtitle of the thesis. The term *translation* is most commonly used to talk about translation from one language to another, but when meaning is transported from one semiotic system to another, for example, from pictures to linguistic form or from books to film, this involves the notion of *intersemiotic translation* (Jakobson, 1959), *transduction* (Kress, 2010), or *intersemiotic transposition* (Sonesson, 2014).

Translation, both within and between semiotic systems, according to Sonesson (2014) is a much more complicated process than merely substituting one code for another. He characterises translation in terms of "a double act of communication" where the translator is "a doubly acting subject, as interpreter and as creator of a new text" (ibid: 263-264). That is, the translator first interprets the artefact according to her/his own understanding, and then makes a choice of either

adapting to the sender of "the first act of communication" (i.e. to the creator of the artefact) or to the receiver of "the second act of communication" (i.e. to the audience), or else compromises maintaining certain aspects of the original work, but adjusting some aspects to fit better in the receiver's background (ibid: 263). Even in regards to translation from one language to another, the meaning can never be identically transferred in the process, but due to system specific features and affordances the content in intersemiotic translation/transposition faces even greater challenges. Thereby, in translation meaning can only be "transferred" within the limits of the semiotic systems involved (ibid).

Each semiotic system has, as we saw in the preceding sections, typical semiotic characteristics and sensory modalities making intersemiotic translation a highly demanding and creative process. For example, language has words, but pictures do not, so the translator is required to find means to transfer the "temporally dynamic output" into "spatially static visual sequences", or vice versa (Berman & Slobin, 1994: 41). Sonesson (2014) argues that the process of transferring meaning from the semiotic system of language into that of pictures faces more challenges than what is true for the process the other way around, since language only has a limited set of qualities at disposal to be "abstracted from the wholes" of perceptual reality (ibid: 274). That is not to say that the pictorial system would be able to deliver everything from perceptual experience, but that it reaches much closer to the perceptual world than language is ever capable of. Therefore, when translating a story from language to pictures, there may not be enough information in the source story available for the adequate transference of meaning, and many new details may need to be added, such as, the way protagonists in the story look, the distance between them, the size and colour of entities, etc. Translating in the opposite direction faces other challenges, for it requires decisions on how to organise the narrative in terms of thematic hierarchies, and also, one has to find the right words to describe the specific aspects of the meaning in the picture (Kress, 2010; Sonesson, 2014). Also, many elements represented iconically would be lost. Thus, when translating from pictures to language, co-speech gestures could be expected to help fill the gap.

Rimé and Schiaratura (1991) report evidence that iconic gestures are likely to be elicited when visual, motoric, or spatial information is translated into speech. We can readily make sense of this finding in terms of our concepts: as gestures have a degree of iconicity that is intermediate compared to the semiotic systems of language (even in the spoken mode) and pictures, when translating from the latter to former, they can to some degree "fill in" for what language lacks.

2.6. Summary

This long chapter offered the theoretical framework for this thesis and introduced human communication as inherently polysemiotic (in terms of semiotic systems) and multimodal (in terms of sensory modalities) in character. Language as a sociocultural phenomenon is grounded in intersubjectivity, and bodily mimesis was offered as a possible explanation for the roots of not only language, but also gestures (and possibly also pictures), and for the close interactions between these semiotic systems. Section 2.2 introduced the field of cognitive semiotics and motivated the use of methodological triangulation in research that aims in understanding "meaning, mind, and communication" (Zlatev, Sonesson & Konderak, 2016). In Section 2.3 the distinct semiotic systems of language, gestures, and pictures were defined using concepts from cognitive semiotics, and description of the polysemiotic nature of human communication finished the section. In Section 2.4 the concepts of narrativity, narratives, and stories were defined as three dimensions of the same phenomenon; a number of narrative structures were defined, and different narrative modes corresponding to the semiotic systems explained. Finally, the notion of (intersemiotic) translation was taken up to present some of the many challenges that need to be met when stories are translated from one semiotic system to another.

2.7. Aims and general hypotheses

We may now reformulate the questions of the present thesis, given in general terms in Chapter 1, with the help of some of the concepts that have been defined in this chapter. We may have the same story, but when it is narrated using two different semiotic systems – speech and a sequence of pictures – we will necessarily have two quite different narratives. The first question concerns how the narrative mode of each semiotic system may influence the respective narrative (possibly including the story content). A second question arises when such pictorial and verbal narratives need to be translated into polysemiotic and multimodal embodied narratives, using speech and gestures: would translation from the narrative mode of speech lead to more *coherent* embodied narratives, and would translation from the narrative mode of pictures lead to more perceptually detailed embodied narratives, as may be expected given the affordances of the "source" systems and their modes? Based on the theory and research described in this chapter, the following general hypotheses can be formulated:

1) Embodied narratives produced as translations from narratives in language will have more narrative coherence, reflected in aspects such as a higher number of plot elements, and more connective devices.

2) Embodied narratives produced as translations from narratives in pictures will have more iconic

representations in both speech (ideophones) and gestures, and above all, enacting gestures (i.e. *acting* in Müller's terms, and *character viewpoint* in McNeill's terms).

Should such differences be found, this could be explained as due to the characteristics of the source systems, which differ from one another not only in perceptual modality but also in regards to expressive potentials and restrictions of the respective semiotic systems.

CHAPTER 3 METHOD

This chapter describes how the experimental data for this thesis was gathered, coded, and analysed following the cognitive semiotic framework with methodological triangulation explained in chapter 2.

3.1. DESIGN

An experiment was designed that consisted of two parts: participants in two groups were first presented the same story in pictures (in pictorial narrative mode) or language (in speech narrative mode), and they then re-narrated the story to an interlocutor. Thus, the embodied narratives elicited in this experiment are based on the same story although they were presented in two different semiotic systems and sensory modalities.

3.1.1. MATERIALS

The embodied narratives for *picture narrative mode* (PNM) were elicited with a sequence of 24 pictures that each represented one or more events and together constituted a story. For *speech narrative mode* (SNM) the embodied narratives were elicited with the help of an oral version of the same story.

The material used in the group of PNM was Mercer Mayer's (1969) picture book *Frog, Where are You?* that was introduced in Section 2.4.3. In this study, however, colour was added to the original black-and-white pictorial narrative to enhance the story experience (see Appendix A). That way the main characters and events became more easily foregrounded, in a similar vein that the semiotic system of speech affords foregrounding with the act of naming and prosodic variation (see Section 2.4.3). Thus, the picture "reading" was made more comparable to that in the speech narrative mode.

The narrative used in the group of SNM was a pre-recorded retelling of the frog story. A script of *Frog, Where are You?* (1969) in English (see Appendix B) was downloaded from *Systematic Analysis of Language Transcripts* (SALT, 2015) webpage, and translated into Finnish specifically for this experiment for the reason that there existed no known written or audio recorded narrative of the publication in Finnish. The translated narrative (see Appendix C) was then recorded as an audio file by a professional voice actor making sure it contained expressive prosody in a way that fits the genre of storytelling.

3.1.2. PARTICIPANTS

Thirty-eight native Finnish L1 speakers (25 females; 13 males), aged 20 to 53 years (mean 34) took part in the experiment in return for a cinema ticket. They were recruited to the study via social media and personal acquaintances. The participants were divided evenly in PNM and SNM groups.

The following factors were controlled between the groups to ensure comparability: the level of education (longer than 12 years), and gender division (6 males in PNM; 7 males in SNM).

3.1.3. Ethical considerations

The participants signed a form of informed consent before the experiment, which included permission to be video-recorded. It was made clear that no personal narratives would be collected in the study and that the participants had the right to withdraw at any time if they so wished. Furthermore, it was stressed that the partakers would remain anonymous. Thus, the video recordings and feedback forms were filed under codenames, such as, *A1* (audio 1), and *V1* (visual 1), to avoid having names documented in connection with the material. The focus of the study was revealed at end of the experiment after the participant had filled in the feedback form (see Appendix D). The participants were also given a chance to freely discuss participant experience and ask questions, which sometimes gave insight into their performance. These discussions, however, were not recorded.

3.1.4. Setup

The experiments were conducted in Finland, in the cities of Helsinki and Oulu, between February and April, 2016. For the sake of ecological validity, naturalness, and interaction, there were a number of factors that were taken into account. Firstly, the location was selected so that possible outside disturbances could be kept minimal (library, home-environment, or quiet café). Secondly, the aim was to make the discourse feeling as natural as possible between the addressee (i.e. myself or a research assistant) and the narrator. Therefore, the participant was an acquaintance (i.e. a friend, a classmate, or a neighbour) of either myself or the research assistant.

3.1.5. Procedure

Participants were divided into one of the two conditions: either the PNM or the SNM. They were told that the study compared reproductions of two different ways to understand a story – in speech or in pictures. They were advised to attend to the narrative carefully, because they would have to retell it to the addressee in their own words right after the exposure. Each participant was exposed to the narratives only once. The length of the audio file was 3 minutes and 30 seconds, and in order to make the length of the narratives identical between the groups, the sequence of pictures was shown in the same pace as the audio recording: the researcher listened to the audio recording on headphones and controlled the page turning of the picture book for the participant. There was only one picture (event scene) per folding. The method of the controlled page turning differs from the "normal" narrative elicitation procedures based on the Frog story where the participants can use the pictures as props for their narratives by freely turning the pages back and forth at their own tempo

when narrating. This alternative method was necessary in order to make the narratives in the two systems comparable. Once the first part of the experiment was over – that is, hearing or seeing the story – the participants were asked to move on to reproducing the narrative to the addressee. After the elicited narrative task, the participants were asked to fill in the feedback form where they were asked to provide some background information concerning, gender, age, handedness, language skills, and possible time lived abroad; and feedback concerning how the experiment had been experienced – that is, whether the task had been challenging, and whether it could have been easier or more difficult had the participant experienced the story in the alternative narrative mode.

3.2. Speech annotation and coding

The reproduced narratives were video-recorded, and later transcribed and annotated in detail using the multimedia annotator ELAN 4.9.1-b, which is a professional software for the creation of multiple layer annotations on video and audio resource (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). When repetition occurred as a result of disfluency, the resulting speech (or speech-gesture combination) was coded once only. Furthermore, incomplete words, interruptions and "pause fillers" (e.g. *mmm, hmm, öööö*) and false starts were deleted.

3.2.1. Clauses

The basic unit of analysis was the clause, defined by Berman and Slobin (1994: 657) as "any unit that contains a unified predicate", that is, "a predicate that expresses a single situation (activity, event, or state)". Clauses classified this way consist of only one verbal element[17], as in (8), or two as in (9). Likewise, embedded clauses were treated as separate clauses, as in (10).

(8) *hyppäsi ulos ikkunasta*
'jumped out of the window'

(9) *poika pyysi koiraa | olemaan ihan hiljaa*
'he asked the dog | to be quiet'

(10) *ja ne meni { sen puun jossa se mehiläispesä oli kiinni } sen alle*
'and they went { the tree where the beehive was attached to } under it'

The clauses were divided into *narrative, metanarrative,* and *extranarrative* units. What is classified as a unit of the narrative kind contains information that is part of the story (Mushin, 2001), possibly an inference based on what the voice actor had narrated on the recording or what had been inferred from the pictures, as in (11) and (12).

[17] Note: following the example of Berman and Slobin (1994: 657) "clauses will be comprised of a single verbal element; however, infinitives and participles which function as complements of modal or aspectual verbs are included with the matrix verb as single clauses, e.g. *want to go, started walking.*"

(11) *ja sit se sammakko karkas*
'and then the frog escaped'

(12) *no siitähän ne mehiläiset suuttu tietenkin*
'of course the bees went mad because of that'

A unit of the metanarrative kind, then, represents a personal stance (e.g. evaluations of the perceived narrative), or the experiences of "the reteller as a reteller in the current context and as a former hearer of the story" (Mushin, 2001: 100)[18], like in (13) and (14).

(13) *mutta mitä ihmettä sitten sen jälkeen tapahtu?*
'but then what happened after that?'

(14) *mut sitä mä en ihan ymmärtäny*
'but that I didn't quite understand'

Extranarrative clauses on the other hand make a comment upon the stimuli or the experiment situation, as in (15) and (16).

(15) *sitä kuva ei paljastanu*
'the picture didn't reveal that'

(16) *siitä puuttu sellanen kohta missä…*
'it was missing an episode where…'

3.2.2. Plot elements

Narrative clauses were divided in different plot components according to the coding scheme provided by Berman and Slobin (1994) introduced in Section 2.4.2:

1) *the onset of the plot* (i.e. the boy realising the disappearance of the frog)
2) *unfolding of the plot* (i.e. the search for the missing frog)
3) *the resolution of the plot* (i.e. the boy finds the frog)

This coding scheme was modified to take into account the whole story from introduction to ending (which were left out in the original coding scheme). Since introduction of the topic and ending of the story are important parts of the narrative structure in verbal narratives, conventional storytelling expressions like *once upon a time there was* were kept and coded as narrative units in the modified scheme. Following this and the clause unit distinctions described in the previous paragraph, the clauses were divided into *introduction, onset of the plot, unfolding of the plot, resolution of the plot*, *ending*, and in *metanarrative*, and *extranarrative* units. This division allowed the narrative and non-

[18] Mushin (2001), however, does not make a distinction between extranarrative and metanarrative clauses, and he solely uses the term *extranarrative* to refer to these narrative units.

narrative units to be analysed separately.

3.2.3. Connecting devices

Clause initial connective devices were identified and counted in each group following the scheme provided by Halliday & Hasan (see 2.4.3). The connective devices were analysed individually, because often they occurred as strings of conjunctions, and they were categorised based on the function of the particular combination. It was found that the function of a connective device in a clause was not always stable, and depended on the context. As can be seen in Table 2, the conjunction *että* 'that'[19] appears in two different categories.

Table 2. List of Finnish clause initial connective devices found in the data and divided in five categories of conjunctive relation based on their function in the clause.

Additive	Adversative	Causal	Temporal	Continuatives
eikä - 'nor'	**että mutta** - 'but that'	**että** - 'that	**eli ensin** - 'so first'	**että** - 'that'
eli - 'that is'	**ja että** - 'and that'	**että kun** - 'that when'	**ja aluksi** - 'and at start'	**ja no** - 'and well'
ja - 'and'	**ja mutta** - 'and but'	**että sitten** - 'that then'	**ja edelleen** - 'and still'	**ja sitten no** 'and then well'
ja kaiken lisäksi 'and above all'	**ja vaikka** 'and although'	**ja että kun** - 'and when'	**ja ensin** - 'and first'	**kuitenkin** - 'anyway'
niin eli 'so that is'	**mutta** - 'but'	**ja koska** - 'and because'	**ja loppujen lopuksi** 'and in the end'	**mutta joka tapauksessa** 'but in any case'
tai - 'or'	**mutta että** - 'but that'	**ja kun** - 'and when'	**ja lopulta** - 'and at last'	**mutta kuitenkin** 'but anyway'
tai sitten 'or then'	**mutta sitten** 'but then'	**ja niin** - 'and so'	**ja no sitten** - 'and well then'	**mutta kuitenkin että** 'but anyway that'
vai - 'or'	**vaan** - 'but	**ja sitten koska** 'and then because'	**ja samaan aikaan** 'and at the same time'	**no** - 'well'
		ja sitten kun 'and then when'	**ja samaan aikaan kun** 'and at the same time that'	**no kuitenkin** 'well anyway'
		ja sitten sen takia 'and then because of that'	**ja samalla** 'and meanwhile'	
		ja sitten siitä syystä 'and then for that reason'	**ja samalla kun** - 'and while'	
		joten - 'so'	**ja sen jälkeen** 'and after that'	
		koska - 'because'	**ja siinä samassa** 'and in that instance'	
		koska sitten 'because then'	**ja siinä vaiheessa** 'and at that point'	
		kun - 'when'	**ja sillä aikaa** 'and meanwhile'	
		mutta sitten kun 'but then when'	**ja sillä välin sitten** 'and meanwhile then'	
		niin että - 'so that'	**ja sitten** - 'and then'	

[19] *Että* conjunctions that introduced direct speech were placed in *continuatives* and those that were found to function as causal connectives were placed in *causal* category.

			ja sitten samaan aikaan kun - 'and then while'	
			ja sitten sen jälkeen 'and then after that'	
			ja tän jälkeen sitten 'and after this then'	
			jonka jälkeen 'after which'	
			koska sitten siinä samalla sitten 'because meanwhile then'	
			kunnes - 'until'	
			niin - 'so'	
			niin ensimmäisenä 'so first'	
			niin silloin - 'so then'	
			niin sitten 'so then'	
			no sitten - 'well then'	
			no sitten kun - 'well when'	
			sen jälkeen - 'after that'	
			siinä välissä - 'at that point'	
			sillä aikaa - 'meanwhile'	
			sillä välin - 'meantime'	
			sitten - 'then'	
			sitten kun - 'then when'	

3.2.4. Story characters

In regards to reference tracking, mentioning of the animate protagonists (*the boy, the dog, the frog, the deer, the gopher, the owl, the bees, the baby frogs,* and *the other frog*) were counted, which made it possible to see which protagonists were mentioned and to compute token per type ratios. All different types of anaphora were taken into account (i.e. full lexical noun, and use of pronominals and null elements).

3.2.5. Ideophones

Ideophones were identified and marked based on two criteria:

- there is an iconic ground between expression and content (see section 2.3.1.)
- an alternative, more or less synonymous, non-iconic expression exists in the lexicon

Interjections, such as *ohhoh,* were not included in the count due to their often indexical rather than iconic ground. Ideophones were further categorised based on their type in regards to modality they map on to (see Table 3 for ideophone inclusion):

- unimodal (i.e. sound to sound resemblance)
- crossmodal (e.g. movement to sound correspondence)

Table 3. Examples for unimodal and crossmodal ideophone, compared to non-motivated "neutral words".

Neutral word	Unimodal Ideophone	Crossmodal ideophone
HUUTAA 'to shout' LENTÄÄ 'to fly'	HUHUILLA 'to hoot'	SUHAHTAA 'to swoosh'

3.2.6. Gestures

Gestures were counted based on strokes, as defined in Kendon (2004). Only gestures that were linked to narrative clauses were considered. Following the definitions provided in Section 2.3.2, gestures were distinguished between *iconic gestures*, *deictic gestures*, *emblematic gestures*, and *pragmatic gestures*. Iconic gestures were divided further in two categories in respect to distinction of whether the articulators were used to display the movement by *enacting* (i.e. first person perspective), or whether they we "non-enacting" in kind (i.e. "symbolic" gestures) where the articulators were used to mold[20] a shape or draw the outline of the object, or to represent an object or movement (e.g. when a hand represents a frog or moving fingers represent moving legs).

3.3. Specific hypotheses

Based on the theory and research described in the previous chapter, and the two general hypotheses formulated in Section 2.7, specific hypotheses can be formulated. Hypotheses 1-4 are operationalisations of the first general hypothesis, and the H5-H8 are operationalisations of the second general hypothesis:

H1) There will be a higher number of plot elements when translated from SNM than PNM narratives due to linguistic cohesion in the source narrative, which provides the foundation for the (linear) unfolding of the narrative plot.

H2) There will be a more diverse use of connective devices (in terms of function) when translated from SNM compared to PNM due to their presence in the speech source narrative, whereas in the picture source narrative the relations between successive events need to be inferred.

H3) There will be a higher ratio of main characters to peripheral characters in the SNM condition compared to PNM; that is, the main characters will be mentioned more often than the peripheral characters. This is expected, because in the SNM condition the three main characters (the boy, the dog, the frog) are "carried" through the narrative through multiple mentions, but

[20] Although "molding" may be considered belonging to the enacting category (Zlatev, 2015b), in this thesis it is regarded as enactment only when performed as imitation.

the several peripheral characters are mentioned only fleetingly without contributing much to the resolution of the plot. In the picture source narrative, however, one of the main characters (the frog) is absent for the most part of the narrative.

H4) There will be more formulaic expressions characteristic to the genre of storytelling in the SNM compared to the PNM condition due to their existence in the source narrative.

H5) There will be a higher number of ideophones in the embodied narratives translated from PNM than from SNM narratives, due to greater presence of iconicity in pictures than language.

H6) The proportion of unimodal ideophones will be higher in the PNM condition when compared to SNM, since unimodal sound symbolism is more transparent (and thus closer to primary iconicity) than cross-modal sound symbolism (which is closer to secondary iconicity).

H7) There will be a higher number of iconic gestures when translating from PNM than SNM narratives, reflecting the more iconic nature of the source narrative.

H8) There will be more enacting gestures when translating from PNM compared to SNM narratives, as these are the kind that most closely correspond to the primary iconicity of pictures in the source narrative.

3.4. Summary

This chapter provided information for how the experiment conducted for this thesis was planned and carried through, and explained the tools and methods that were used for data analysis. The chapter ended with detailed hypotheses that were operationalisations of the general hypotheses provided in Chapter 2, which on their side were derived from the research questions given in Chapter 1.

CHAPTER 4 RESULTS AND DISCUSSION

4.1. INTRODUCTION

This chapter presents the results of the research and links them to the detailed hypotheses provided at the end of Chapter 3. The chapter is divided in three main sections: the two first present the results of which speech analysis is given in 4.2, and gesture analysis in 4.3. General discussion is provided in 4.4 followed by a short summary in 4.5.

The statistical results presented in this chapter are based on Poisson regression (for rates) and logistic regression (for proportions). The results are the estimate of the *effect* (EST), its *standard error* (SE), the z-value and the p-value. The main predictor was always semiotic system (i.e., PNM or SNM). In some of the analyses, clause length was included as a covariate. This was done, for instance, in the analysis of the ideophone rates. The effect of clause length is not of interest to the hypotheses and is therefore not reported below. In most analyses, finally, participants were also included as a random predictor.

The results of the statistical analyses for all hypotheses are collected in Table 4 for a quick review, and the descriptive results related to each of these hypotheses are collected in Table 5 together with proportional statistics and occurrence rates when applicable.

Table 4. Statistical difference between PNM and SNM in H1-H8. Where the difference is not statistically significant, the area has been marked grey. The results are the estimate of the effect (EST), its standard error (SE), the z-value, and the p-value.

	EST	SE	z	p
H1 Plot elements	-0.799	0.467	-1.712	0.087
H2 Connective devices	-0.193	0.108	-1.782	0.075
additive	0.049	0.180	0.274	0.784
adversative	-0.873	0.243	-3.585	0.000
causal	0.286	0.239	1.197	0.231
temporal	-0.047	0.219	-0.216	0.829
continuatives				
H3 Ratio main to peripheral characters	-0.150	0.088	-1.717	0.086
H4 Formulaic story expressions	0.005	0.042	0.126	0.900
H5 Ideophones	-0.023	0.043	-0.532	0.595
H6 Ideophone modality	0.161	0.395	0.407	0.684
H7 Iconic gestures	-0.199	0.366	-0.543	0.259
H8 Enacting gestures	2.303	0.599	3.846	0.000

Table 5. Overview of the results related to H1-H8.

			Proportions		Rate per clause (per word)	
	PNM	SNM	PNM	SNM	PNM	SNM
Narratives	**19**	**19**				
Words	7848	6822				
Clauses	1285	1117				
Average clause length (words)	6.10	6.10				
H1 Narrative structure / plot elements (clauses)	**1237**	**1042**	**0.97**	**0.93**		
introduction	136	61	0.11	0.06		
onset of the plot	86	72	0.07	0.07		
unfolding of the plot	895	763	0.72	0.73		
resolution of the plot	109	141	0.09	0.14		
ending	11	5	0.01	0.00		
Metanarrative clauses	**42**	**66**	**0.03**	**0.06**		
Extranarrative clauses	**6**	**9**	**0.00**	**0.01**		
H2 Connective devices	**931**	**719**			**0.69**	**0.75**
additive	433	318	0.46	0.44		
adversative	42	75	0.05	0.10		
causal	84	46	0.09	0.06		
temporal	295	253	0.32	0.35		
continuatives	77	27	0.08	0.04		
H3 - Story participants	**1826**	**1543**				
main characters	1471	1205	0.81	0.78	1.15	1.08
peripheral characters	355	338	0.19	0.22	0.28	0.30
H4 - Formulaic story expressions	**16**	**19**			**0.013**	**0.018**
H5 - H6 Ideophones	**95**	**56**			**0.077**	**0.054**
unimodal	74	42	0.78	0.75		
crossmodal	21	14	0.22	0.25		
H7 - H8 Gestures	**321**	**534**			**0.259 (0.04)**	**0.512 (0.08)**
deictic	3	6	0.01	0.01		
emblematic	14	4	0.04	0.01		
iconic	229	418	0.71	0.78		
- enacting gestures	46	10	0.20	0.02		
- non-enacting gestures	183	408	0.80	0.98		
pragmatic	75	106	0.23	0.20		

4.2. Speech analysis

The first six of the eight hypotheses formulated in the previous chapter relate to language, and in this section they are linked to the findings in consecutive order, starting from H1 and ending in H6.

4.2.1. H1 – Plot elements

As can be seen in Table 5, the embodied narratives derived from the PNM condition were longer compared to those produced in the SNM condition. Although there were no differences in the average clause length between the conditions, the participants who had seen the story in pictures

produced in total almost 200 narrative clauses more than the participants who had heard the story. Proportionally the difference in the number of plot elements between PNM and SNM is marginally significant (see Table 4); however, the difference goes in the opposite direction than expected (see Figures 8 and 9). Thus, H1 was not supported. Interestingly, there were more metanarrative and extranarrative clauses in the SNM condition (see Figure 9).

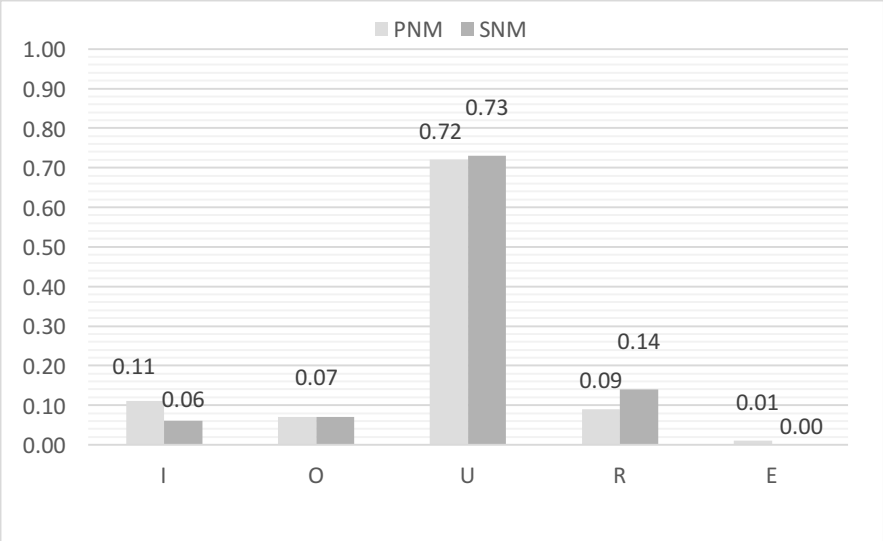

Figure 8. Proportions of narrative clauses: introduction (I), onset of the plot (O), unfolding of the plot (U), resolution of the plot (R), and ending (E).

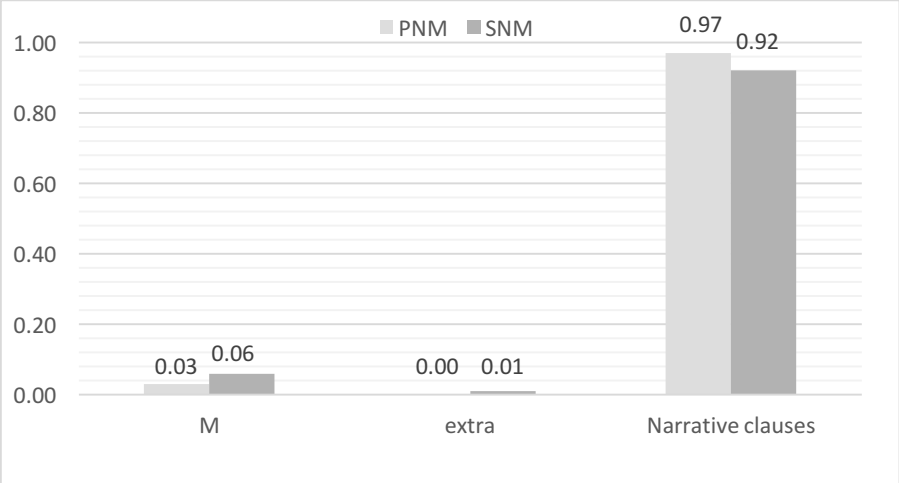

Figure 9. Proportions of metanarrative (M) clauses, extranarrative (extra) clauses, and narrative clauses.

4.2.2. H2 – Connective devices

Overall there were fairly higher rates of connective devices in SNM than in PNM, a difference that is marginally significant (see Table 4). The most frequently used connective devices in both conditions were in the *additive* category, followed by the *temporal* category, as can be seen in

Figure 10. Statistically the use of connectives in those two categories and in the *causal* category did not differ significantly. However, the connective devices in the *adversative* category were used twice as often in SNM than in PNM, whereas the opposite was true for the connectives in the category of *continuatives*. The difference in these latter two categories was significant between PNM and SNM (see Table 4). Thus, while H2 cannot be said to be supported as the results do not show a more "diverse" use of connective devices in terms of function in SNM compared to PNM, the differences in adversatives and continuatives could perhaps be explained by the source narrative mode. We will return to this in Section 4.4.

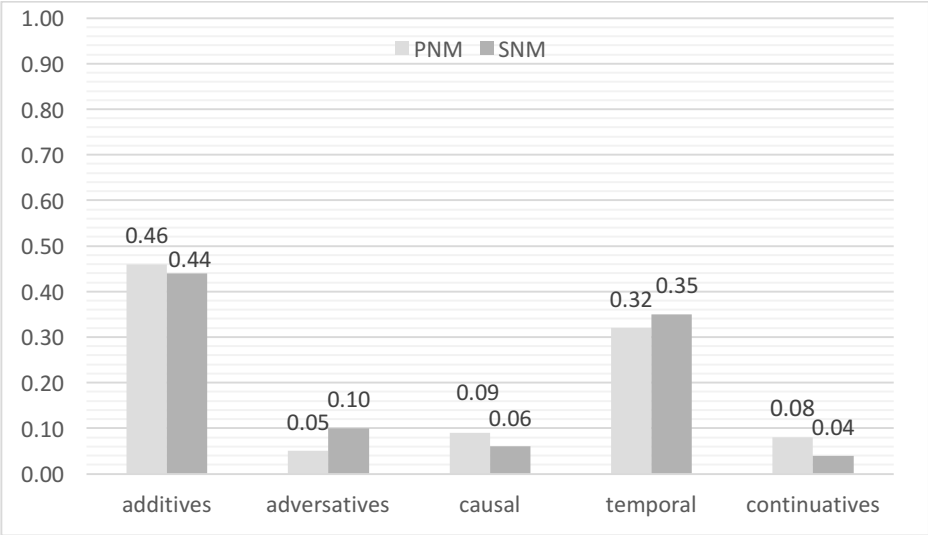

Figure 10. Proportions of connective devices in the five different categories of conjunctive relation.

4.2.3. H3 – Story participants

In both of the conditions the main characters of the story (*the boy*, *the dog*, and *the frog*) got more mentions than the peripheral characters (*the bees, the deer, the gopher, the baby frogs, the partner frog*, and *the owl*) (see Figure 11). As can be seen in Table 4, there was a marginally significant difference in the ratio of main characters to peripheral characters between PNM and SNM; however, as can be seen in Figure 11, this difference was in the opposite direction to what was expected. Thus, H3 was not supported.

4.2.4. H4 – Formulaic expressions from the genre of storytelling

Two percent of narrative clauses in SNM were identified as formulaic story expressions, and in PNM the proportion was one percent (see Figure 12). The frequency of these expressions thus is higher in the group that heard the story, as expected. The difference, however, was not significant between the conditions, as can be seen in Table 4. Thus, H4 was not supported as well.

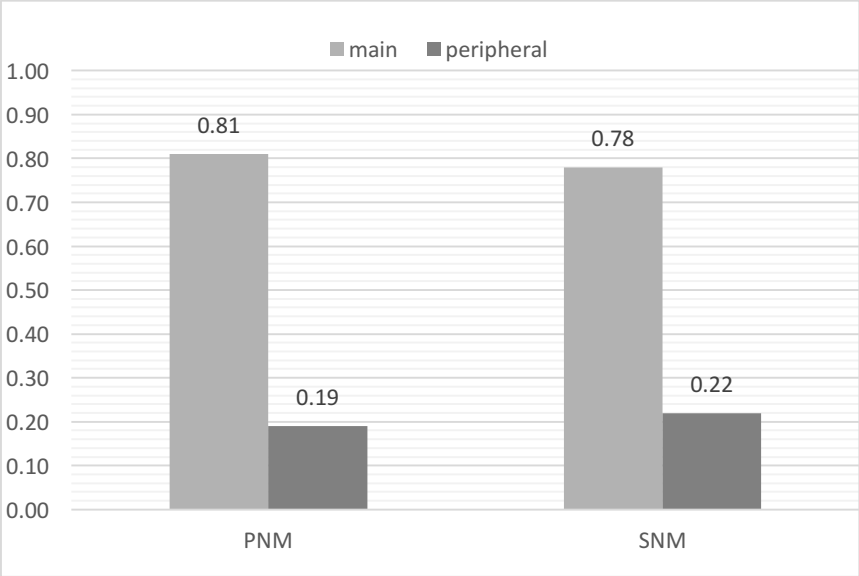

Figure11. Proportional ratio of main and peripheral characters.

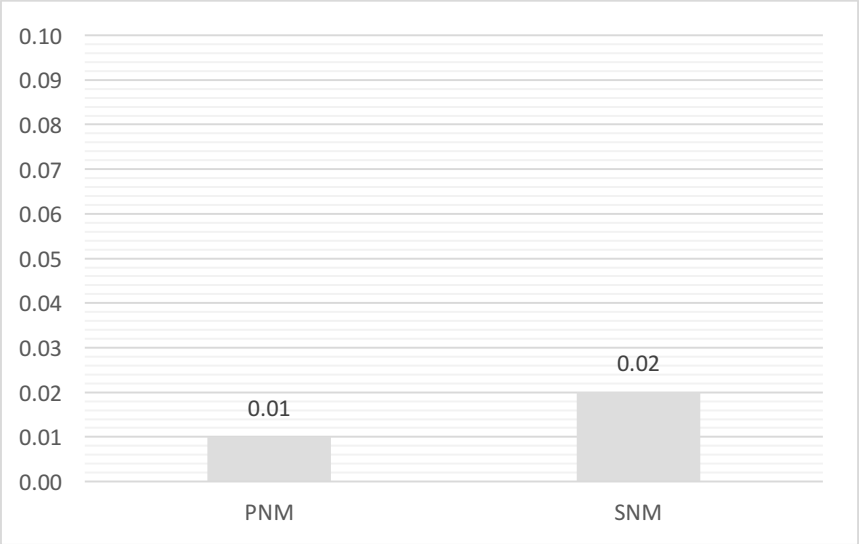

Figure 12. Proportion of formulaic story expressions.

4.2.5. H5 – IDEOPHONES

There were 151 ideophone tokens in the data, and as expected, their rate per clause was higher in the group of participants who had experienced the story in pictures rather than in the group who had heard the story (see Figure 13). However, the difference between PNM and SNM was not statistically significant, as can be seen in Table 4. Thus, H5 does not receive significant support.

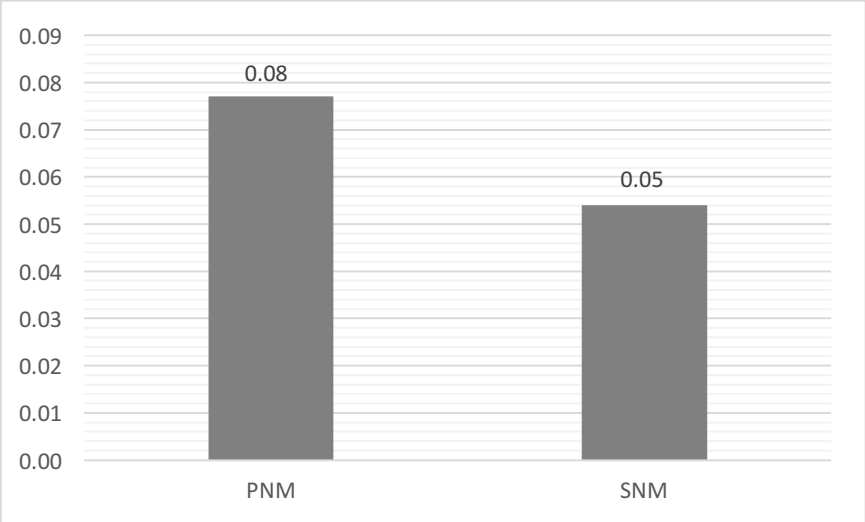

Figure 13. Ideophone rate per narrative clause.

4.2.6. H6 – IDEOPHONE MODALITY

The total number of ideophones consisted of 26 types (i.e. lexical roots). These ideophones were categorised based on the perceptual modality they mapped onto – and thus whether the iconicity in the word was unimodal (i.e. sound to sound resemblance) or crossmodal (e.g. movement to sound correspondence). It was hypothesised that the proportion of unimodal ideophones would be higher in PNM and thus that of crossmodal ideophones higher in SNM. The results show the expected (see Figure 14), however, the difference between the groups was not statistically significant (see Table 4). Thus, H6 was not supported.

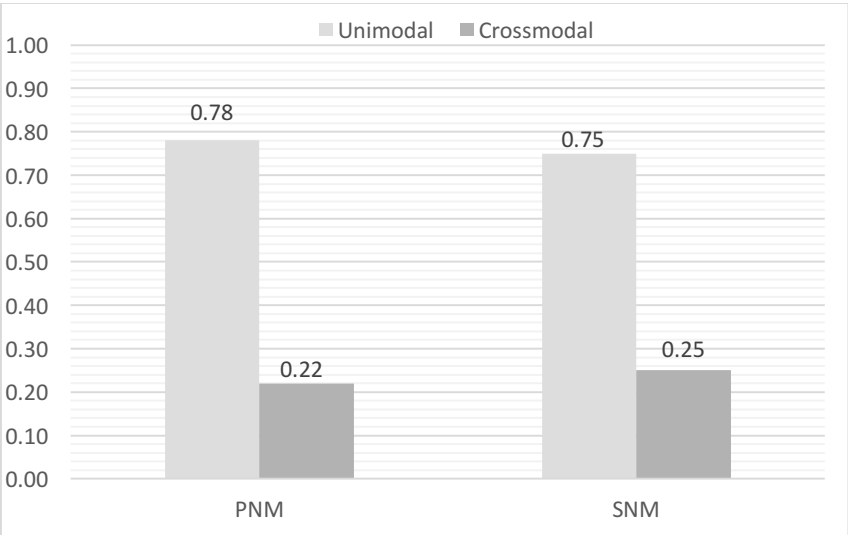

Figure 14. Proportions of ideophone modality.

4.3. Gesture analysis

The last two of the eight hypotheses formulated in the previous chapter relate to gestures, and in this section H7 and H8 are linked to the findings of the study.

4.3.1 H7 – Iconic gestures

As can be seen in Figure 15, the frequency of gestures in SNM was twice the frequency of gestures in PNM. Iconic gestures were proportionally the largest group in both conditions when dividing gestures in different types according to their main function (deictic, emblematic, iconic, or pragmatic) (see Figure 16). Although there is a substantial difference in the frequency of iconic gestures between PNM and SNM conditions (see Table 6), when taking into account the number of narrative clauses in each condition, the difference does not show to be statistically significant, as Table 4 shows. Furthermore, the number of iconic gestures in SNM is counter to our expectations, for PNM was expected to result in a higher proportion in their use, not SNM. Thus, H7 was not supported. In the following section we will return to discuss what might have resulted in the higher use of gestures in the SNM compared to the PNM condition.

Figure 15. Gesture rate per clause and per word.

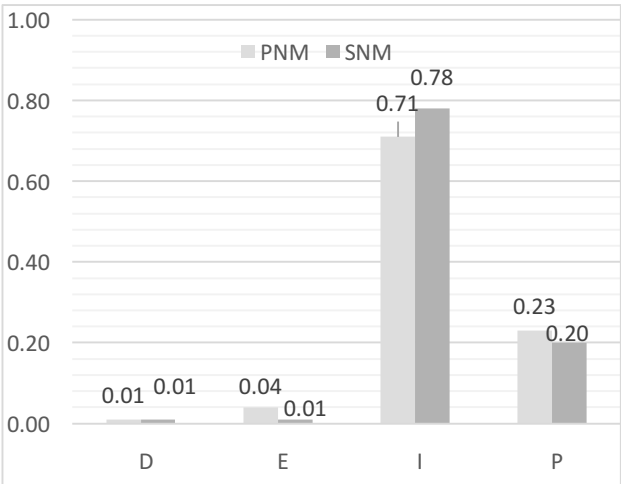

Figure 16. Proportional division of gestures in functions: deictic (D), emblematic (E), iconic (I), and pragmatic (P).

Table 6. Proportions of iconic gestures counted from the total number of iconic gestures across conditions (637 gestures); then in each condition when compared to the other gestures in the group (321 in PNM and 534 in SNM); then per narrative clause, and per word.

	PNM	SNM
All iconic gestures	0.35	0.65
All gestures in the condition	0.71	0.78
Per clause	0.19	0.40
Per word	0.03	0.07

4.3.2. H8 – ENACTING GESTURES

Although there were nearly twice as many iconic gestures in the speech condition compared to the picture condition, the frequency of enacting gestures nevertheless was higher in the latter (see Table 5). In PNM 20% of the iconic gestures were enactments (or CVPT), whereas only 2% of the iconic gestures in SNM were these (see Figure 17). This finding is in line with the hypothesis that there will be more enacting gestures in PNM compared to SNM condition. As can be seen in Table 4., the difference between the conditions is statistically significant. Thus, H8 is supported.

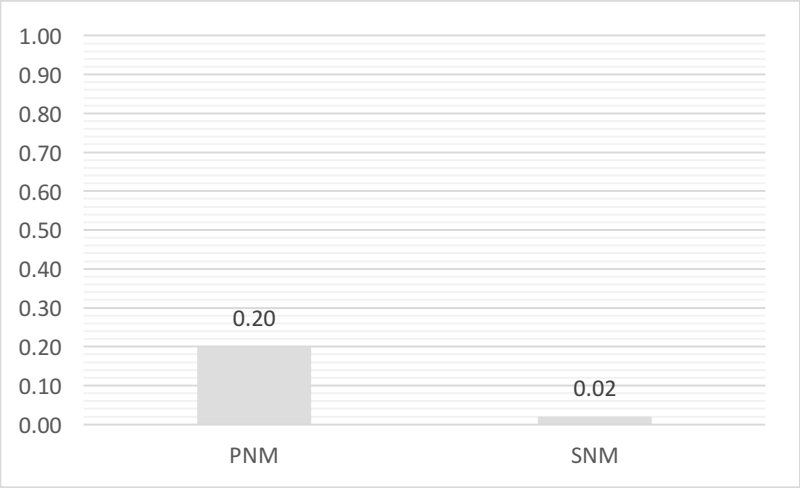

Figure 17. Proportion of enacting gestures counted from iconic gestures.

4.4. DISCUSSION

Although the majority of the hypotheses did not receive support in the study, the results presented in the previous section nevertheless indicate that the sensory modality and/or the semiotic system of the source narrative may indeed lead to differences when translating unimodal narratives into polysemiotic, embodied narratives.

Unlike expected by H1, the proportion of plot elements was not higher in SNM compared to PNM. Concerning H2, while the results did not show "more diverse" use of connectives, the significantly different proportions of adversatives and continuatives in the two conditions do suggest that the source system could lead to some differences in the use of connective devices. H3 was also not supported: the ratio of main characters to peripheral characters was slightly higher in the picture condition, and that of the peripheral characters in the speech condition. Although there was a difference along the lines of the expected direction between the tested conditions in H4, H5, and H6, it was not statistically significant. Counter to the expectations of H7: the number of iconic gestures in SNM was nearly twice the number in PNM – however, most of the iconic gestures in SNM were non-enacting in type, as expected. Finally, the only hypothesis that was clearly

confirmed was H8. This hypothesis presupposed connection of iconicity between the source narrative mode and the embodied narrative through presence of more highly iconic enacting gestures. Because most of the enacting gestures indeed occurred in PNM, the hypothesis was supported.

As noted, the participants who had experienced the story through pictures produced longer narratives than the participants who had experienced the story through hearing. Although language may afford more freedom in terms of how one may imagine the perceptual storyworld (e.g. the appearance of the characters) in comparison to pictures where the perceptual details are directly given, the picture narrative mode could still be argued to afford more freedom in another sense. It allows the narrator to improvise with the narrative details in a way allowing for a creation of a more creative story around certain core elements given in the pictures, thus affording a greater semiotic freedom of interpretation. One clear difference in the plot elements between the conditions was found in the number of clauses produced for the introduction to the story. The participants who had heard the story tended to move directly to the onset of the plot after giving a brief introduction of the main characters and possibly of the setting, which was often only a mention of the frog's jar, as in (17) or that the boy and the dog were sleeping. The introductions produced by the participants who had seen the pictorial narrative on the other hand usually contained more detailed descriptions of the setting including the spatial arrangement of the characters and the details of the surroundings, the characters, and descriptions of the activities the characters were engaged in, as in (18).

(17) *olipa kerran poika . pojalla oli lemmikkinä koira ja sammakko . sammakko asu lasipurkissa.*
'once upon a time there was a boy . the boy had a dog and a frog as pets . the frog lived in a glass jar .'

(18) *poika ja koira oli illalla sängyn vieressä . oli kuunsirppi taustalla .*
ja ne katsoi sinne purkkiin . kun siel oli sammakko . joka oli selvästi tämmönen niiden oma lemmikkisammakko . sitten ne meni onnellisesti nukkumaan . sammutti valot .
'a boy and a dog were next to a bed in the evening . a crescent was there on the background . and they were looking into the jar . because there was a frog . that was clearly their own sort of a pet frog . then they went happily to sleep . put the lights off .'

Such dissimilarities could be explained by the differences in the source semiotic systems. As explained in Chapter 2, there are more perceptual details present in pictures than in words and sentences, and framed space is one of the affordances of the semiotic system of pictures that language lacks. Moreover, although – as noted earlier – words can induce different mental images depending on a person (e.g. a prototype of a dog can be a German Shepherd to one and a Jack

Russel to another), scenes in pictures can provide a no less and possibly richer setting for narrative hypothesising giving rise to "theories" and questions depending on the viewer (see 2.4.3). This is related to another factor that explains the higher rates of introductory clauses in PNM: some participants produced an introduction that preceded the first scene in the story, explaining why the story started from the certain settings, as in (19) – that is, how the frog ended up into the jar in the boy's bedroom in the first place. This could be thought as *naturalisation* in terms of Fludernik (1996), as discussed in 2.4.2 – a way to give a rational motivation for the story, since frogs are not the most typical animals to keep as pets.

(19) *olipa kerran Ville ja sen koira Putte . he olivat metsässä sitten olleet . etsimässä erilaisia käpyjä . ja kuin ollakkaan niin oli löytynyt sitten sieltä tommosen puun kolosta niin sammakko. no Villehän päätti sit sen sammakon ottaa . ja pistää lasipurkkiin . ja viedä kotiin . ja tää Putte-koira oli ihan ihmeissään . et miten tää sammakko voi tääl purkissa olla.*
'once upon a time there was Ville and his dog Putte . they had been in the woods . looking for different kinds of cones [strobili] . and how about that in the hollow of a tree they had found a frog. well Ville of course decided to take the frog . and put it in a glass jar . and take it home . and Putte the dog was in a state of utter bewilderment . [wondering] that how can the frog be in this jar .'

In the SNM condition, on the other hand, it was rather side comments that were used as a means to naturalise the target narratives. The participants that had heard the story tried to stay as close to the original story as possible, and by producing metanarrative and extranarrative clauses indicated what was found to be incoherent in the source narrative, as in (20), or indicating flaws in their own production of the story, as in (21). This could explain the higher rates of non-plot elements in this condition.

(20) *ja sit semmonen niinku kummallinen kohta siin sadussa oli siis vaan se . jos mä saan kommentoida . et siin ei ollu mitään semmosta . et poika riensi rappusia alas tai näin . mul meni vähän keskittyminen siinä kohtaa .*

'and then a weird part in that story was just . if I can comment . that there wasn't anything like . that the boy rushed down the stairs or so . my concentration suffered a bit at that point.'

(21) *mä kyl oon unohtanu jonkun vaiheen tässä .*
'I have forgotten some part here now .'

When evaluating the "fluidity" of the narrative performance, the narratives perceived visually seemed to have been internalised to a higher degree than those perceived by hearing. The participants in SNM more often seemed to be restricted by the source narrative system in their own retelling performance, which could be seen, for example, in uncertainty in the choice of words, in the order of the events and in some particularities. Such restrictions could lead to shorter narratives and explains further the higher rates of metanarrative clauses, such as (22) and (23).

(22) *ja en muista miten se liitty tähän tarinaan .*
'and I can't remember how it's related to this story'

(23) *[se sammakko löyty jostaki .] mut en muista mistä .*
'[from somewhere the frog was found .] but I can't remember where .'

The reason for this difference between the conditions could be that language is more descriptive and characterised by the sequential arrangement of elements in time, whereas pictures are more demonstrative and mimetic, and the information in a picture is available simultaneously (see 2.4.3). Thus, although the pacing of the scenes was controlled and equal between the conditions, the unfolding of information between them differed. This seemed to have an effect on how the story was internalised in each condition. For the reason that the semiotic system of pictures affords a greater freedom of interpretation and therefore does not restrict the narrator in terms of "accuracy" in the narrative production as much as the semiotic system of language, the embodied narratives in PNM appeared more fluid and creative than those in SNM condition. Although not statistically significant (EST = 0.025, SE = 0.329, t = 0.077, p = 0.939), the difference between the conditions was seen also in the narration times: the PNM participants narrated slightly faster (0.05 minutes per clause) than the participants in SNM (0.06 minutes per clause). Considering that the intersemiotic translations in PNM were longer in terms of words and clauses than those in SNM, and yet in total the two conditions took roughly the same time to narrate (61.78 minutes in PNM; and 61.30 minutes in SNM) supports the view that the embodied narratives in PNM were more fluid. That is, the information was more readily available for the narrators in the picture condition, whereas in the speech condition remembering the story details took more effort, and consequently the retold narratives in SNM appeared less "fluent". This was also confirmed by the participants' evaluation of their own narration performance in the feedback form given at the end of the experiment: 11 out of 19 participants in SNM regarded the task challenging, and 15 believed they would have remembered the story events better had they seen the story in pictures. In the PNM condition only 6 participants out of 19 regarded the task "slightly" challenging, and only 2 believed they could have understood or renarrated the story better if they had heard the story instead of seeing it.

Whereas uncertainty in SNM with word choices seemed to be based on attempts to remember the right words the voice actor had used in the source narrative, as in (24), in the PNM condition the participants also had difficulties with word choices, which in their case was based on identification of the animals represented in the pictures, as in (25). This particular example also shows how picture narratives may get different interpretations, because the participant here refers to the event represented in Figure 18 – usually interpreted as the gopher biting the boy on his nose – as the boy's reaction to bad smell.

(24) *ja se koira kiinnitti huomion johonkin ampiaispesään . vai oliko se mehiläispesä ?*
'and the dog paid attention to some wasps' nest . or was it a beehive ?'

(25) *ja sitten oli joku myyrä tai tämmönen pikkuelukka . joka tuoksu pahalle . oisko se ollu joku haisunäätä tai jotain vastaavaa .*
'and then there was some vole or this kind of a small creature . that smelled bad . maybe it was a skunk or something .'

Figure 18. Picture 10 in the picture source narrative that presents the gopher biting the boy's nose while the dog is harassing the bees (adapted from Mayer, 1969).

Furthermore, due to the differences in the level of internalisation of the story content, and the unfolding of information in the source narrative, it also makes sense that a narrator that narrates based on the speech source narrative might easier confuse who did what to whom, as in (26), while in the corresponding picture of the picture source narrative (see Figure 19), the correct position of the boy is more likely to be remembered. Accuracy in story details was not among the tested

factors, but it nevertheless played a role in the experiment, for as demonstrated above, uncertainty often resulted in hesitation or manifestation of metanarrative or extranarrative clauses.

(26) *joka [peura] sit niinku nappas jotenkin sarviensa väliin sen koiran . ja lähti juoksee . ja sit se poika lähti niinku juoksee sen peuran perässä .*

‘who [deer] then like grabbed the dog somehow between its antlers . and started running . and then the boy like started running after the deer .’

Figure 19. Picture 16 in the picture source narrative (adapted from Mayer, 1969) that corresponds with line 16 in the English transcript (see Appendix B): “The deer started running with the boy still on his head. The dog ran along too. They were getting close to a cliff.”

In the PNM condition it was demonstrated by one participant that if a viewer misses some pictorial information in the beginning of the story, this could lead to a completely different story altogether. In the speech source narrative, the frog is present throughout the narrative, but in the picture narrative condition the frog is present only in pictures 1 and 2, and does not reappear until picture 22, which marks the first scene of plot resolution. Since the participant managed to miss the presence of the frog in the beginning of the picture source narrative, instead of narrating a story of a quest for the missing frog, what was narrated was a story of a boy and dog’s adventures, which is also how the participant named the story (see 27).

(27) *tota pojan ja koiran seikkailut vois olla vaikka tarinan nimi .*

‘well adventures of a boy and a dog could be the name of the story .’

But even without missing information in the narrative, the greater freedom of interpretation afforded by the semiotic system of pictures – compared to that of language – can result in various

different ways of telling of the "same" story, as was explained in 2.4.3. This was also demonstrated in our study, for many of the PNM narrators improvised and freely created a personalised story by naming the characters and giving various motivations for their actions, as several of the excerpts in this section illustrate. In this way in embodied narratives *narrative* can add to the *story*.

The finding that arose in this study as one of the most prevalent characteristics differentiating the two narrative conditions from one another was the level of *experientiality*. The noticeable aspect that could be linked to experientiality in the embodied narratives is the presence of the consciousness of the protagonist or the consciousness of the narrator (see Section 2.4.3). Interestingly, this was found to be more typical in the picture narrative mode than in the other condition, and was manifested in the frequency of direct speech, reflecting character consciousness (i.e. shifting the perspective from third person narration to first person narration), as in (28). There were 74 instances of first person narrative clauses (i.e. direct speech) in the picture condition and only 12 first person clauses in the speech condition, with significantly different proportions (see Figure 20).

(28) *no eihän me nyt sitte voida tollasta niinku isää tai äitii viedä . että mut jotain pitää viedä . niin viedään sitte yks niistä lapsista .*

'now we can't take a mom or a dad . but something we have to take . so let's take one of the kids .'

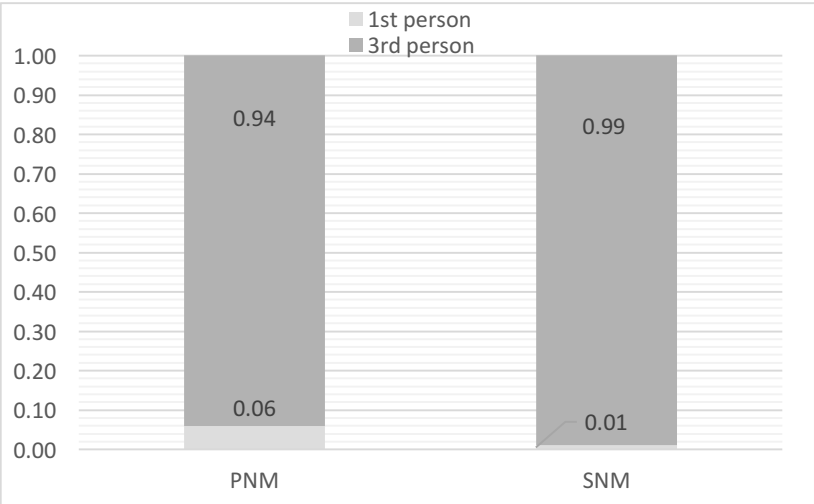

Figure 20. Proportion of first person direct speech and third person indirect speech perspective. The difference between the conditions is significant (EST = -1.573, SE = 0.495, z = -3.176, p < 0.001)

Mostly these instances of direct speech in SNM were those that were also present in the source narrative: the boy calling the frog, as in (29). In PNM the perspective was chosen more freely between any of the main or peripheral characters. Further, the embodied narratives in the PNM condition were also more often found to reflect "experiential value" through empathy, as in (30) and (31).

(29) *missä olet sammakko ?*
'where are you frog ?'

(30) *ja eihän Max kun ei ollu aikasemmin nähny . ni ei tienny . et se oli ampiaispesä .*
'and of course since Max had never seen alike before . didn't know . that it was a beehive.'

(31) *ja pöllökään ei diggaa siitä . ku se tsiigaa . et sen frendi mehiläisten pesä on rikottu .*
'and the owl doesn't like it either . when he sees . that his bee friends' nest has been broken.'

In terms of Fludernik's (1996) continuum of experientiality the embodied narratives in the SNM condition tended to have more properties from the *narrative report* style being built upon "second-hand experience or on a summary of first-hand experience rendered non-experientially" (ibid: 71). As discussed in 2.4.3, this narrative style Fludernik connects with the use of connective device *then*. According to Halliday & Hasan's (1976) scheme of conjunctive relations (see 2.4.3) *then* locates in the category of *temporal* connectives. Although this category indeed was proportionally slightly higher in SNM compared to PNM, the difference was not statistically significant. The embodied narratives in the PNM condition in contrast seemed to have more properties from *narratives of vicarious experience* in terms of Fludernik's continuum. This narrative style was connected with the use of connective *so* (see 2.4.3), which locates in the category of *causals* in the scheme of Halliday & Hasan (1976). Also in our findings this category was indeed proportionally higher in PNM compared to SNM, but like with *then*, the difference between the groups was not statistically significant.

The connective device categories that, however, were found to differ significantly between the tested conditions were *adversatives* (higher in SNM) and *continuatives* (higher in PNM). Although these may not link to the continuum of experientiality, their differences between PNM and SNM may also be connected to the source narrative systems. Firstly, the most frequent adversative connective *mutta* 'but' was in many occasions found to be preceded by a metanarrative or extranarrative clause, as in (32).

(32) *sitä tarina ei kertonut .* ***mutta*** *poika oli onnellinen .*
'the story didn't tell that . **but** the boy was happy .'

The adversatives were often also used to emphasise contrast, as in (33), or surprise, as in (34). Adversative connectives both presenting contrast and expressing that something was counter to expectations were in place in the speech source narrative, and since their occurrence was higher in SNM compared to PNM, we can readily interpret their frequency stemming from the source narrative.

(33) *sille koiralle ei käyny kuinkaan* . ***mutta*** *purkki meni rikki* .
'nothing happened to the dog . **but** the jar broke '

(34) ***mutta*** *oksat eivät olleetkaan oksia* .
'**but** the branches weren't branches after all .'

Secondly, the most frequent continuative conjunction *että* 'that' was often used to introduce direct speech, as in (35), and since the majority of first person perspective narration happened in the PNM condition, this could explain the difference in the continuatives between the conditions.

(35) ***että*** *älä mee sinne* . ***et*** *ampiaiset tulee* . *ja pistää sinua* .
'**that** don't go there . **that** the bees will come . and sting you .'

In regards to formulaic expressions (H4) it could be suggested is that there were no significant statistical differences because although there were story expressions in place in the SNM source narrative, they did not transfer into the embodied narratives as amply as expected for the reason that, in general, the story did not seem to be as well internalised as in the other group. Namely, as mentioned, the narratives in SNM were in many occasions closer to the narrative report style with the main function to provide information more than storytelling. Even though not in place in the source narrative, the higher degree of experientiality in PNM on the other hand resulted in a higher than expected frequency of these storytelling expressions.

Although the embodied narratives in PNM were overall more "vivid" and creative, this nevertheless was not reflected in more frequent use of ideophones as anticipated. Although indeed higher in number in PNM compared to SNM, the difference was not statistically significant. Similarly, statistical differences regarding ideophone modality between the conditions was not found. Nonetheless, one feature worth mentioning is that in connection with direct speech, it was not uncommon for the participants in PNM to use ideophonic interjections, which sometimes were combined with enacting gestures – for example, *shh* or *hys* 'hush' indicating a request to be quiet, or exclaiming *huh* 'phew' that indicated relief. Although most of these interjections in the source narrative were indexical (i.e. they had an associative, indexical ground rather than iconic), in the target narratives they were in fact iconic, because each time of their occurrence they represented sounds made by protagonists.

Table 7. List of interjections in PNM.

INTERJECTIONS		
huh 'phew'	jahaa 'ooh'	shh
hys 'hush'	ohhoh 'oh my'	voi 'oh'
		voi voi 'oh no'

Since the target narratives produced in the SNM condition were shorter in terms of narrative clauses and yet combined with a higher number of gestures and took more time to narrate than the narratives in the PNM condition, a closer inspection of the gestures in connection to speech was required. As we have seen, the demand on memory and the ability to internalise the story has an effect on the level of apparent narrative experientiality. In Section 2.3.4 it was suggested that gesturing may help shifting load from verbal working memory to other cognitive systems or external representations, and the findings of this study could indeed indicate a facilitative role of gestures – namely, that of iconic non-enacting gestures – on narrative production.[21] The listener of a spoken narrative does not have the freedom to construct a personalised story the way a viewer of a picture narrative has, thereupon she/he is required to rely on memory in a different manner in the retelling of the story. As mentioned, in terms of Fludernik's (1996) continuum of experientiality, the embodied narratives in the SNM condition tended to have more properties from the *narrative report* style being built upon "second-hand experience or on a summary of first-hand experience rendered non-experientially" (ibid: 71). Interestingly the gestures of the SNM participants were found to reflect this level of experientiality. The narrators in this condition more often than in PNM narrated the story from the "outside" of the storyworld, and in a way by help of gestures drew a visual map of the story events and characters in front of them. This systematic use of gesture space could be seen as a way to create and maintain cohesion in the embodied narrative, which not only is helpful for the interlocutor, but also for the narrator in the process of narrative production. The embodied narratives in PNM, in contrast, seemed to have more properties from *narratives of vicarious experience* in terms of Fludernik's continuum, which could be seen, for example, in a greater use of first person perspective in both gestures and speech (thus reflecting a high level of experientiality). Instead of drawing a map of events in front of them, the narrators seemed to

[21] *Narrative production* here should not be mistaken with *speech production,* but rather understand it as "story creation".

position themselves *inside* the storyworld. This explains the occurrence of first person enacting gestures (and direct speech) in the picture narrative condition, but also in connection to non-enacting gestures, the gaze was found to be one of the indicators of the narrator's position either inside or outside the storyworld. This outside versus inside the story perspective could be the reason for the substantially different rates of gestures between SNM and PNM (although the difference did not reach statistical significance). Namely, managing the visual map of the storyworld from the outside would require a greater use of gestures than when one positions oneself inside the story, being able to refer to things in relation to one's own body. Correspondingly, narrating a story based on second hand experience is more challenging than when being able to express the events through access to the consciousness of the protagonists – that is, through perceptual focalisations, or through empathy.

As mentioned, direct speech and narration through empathy reflecting the presence of a conscious actor (i.e. manifestation of a higher level of experientiality) was often combined with first person perspective in gestural expression, which is possibly the reason for the higher occurrence of iconic enacting gestures in PNM compared to SNM condition. This evidence links with Rimé and Schiaratura's (1991) finding – discussed in section 2.5 – that iconic gestures are likely to be produced when visual, motoric, or spatial information is translated into speech. However, in light of our findings, this can only clearly be said about iconic *enacting* gestures, not iconic non-enacting gestures (i.e. when articulators are used to mold, draw, or represent content) – for H8 was supported, but H7 was not.

4.5. Summary

This section presented the results and linked them to the hypotheses of the study. Only one hypothesis received clear support, but the findings nevertheless suggest that the semiotic system of the source narrative appeared indeed to have an effect on the embodied narratives. Namely, the semiotic system of language (specifically, its subsystem speech) resulted in shorter embodied narratives in terms of clauses, but a longer narration time and a higher rate of iconic non-enacting gestures, which was seen reflecting a lower level of experientiality in the SNM condition. In effect, the perspective on the story events was more distant, leading the narrators to narrate the story from the outside of the storyworld, in which they got support from a systematic use of co-speech gestures. On the other hand, the semiotic system of pictures reflected iconicity in the retelling performance through presence of more iconic enacting gestures, as expected. The embodied narratives in this condition were characterised by a first person perspective in both gestures and in speech. This finding indicates a high degree of experientiality in the PNM condition, and the

intersemiotic translations in this condition may be considered as more vivid and fluid storytelling compared to SNM, resulting from the embodied narratives taking a perspective inside the storyworld. The vividness of the PNM narratives possibly stems from a greater semiotic freedom afforded by the pictorial system, which was often seen as personalised improvisations where the embodied narrative "added" to the story. The SNM condition instead resulted in more report style narratives, because the semiotic system of language allows less semiotic freedom of interpretation and arguably causes more demand on memory.

When translating from unisemiotic and unimodal narratives realised as either language or pictures to multimodal polysemiotic narratives, strictly speaking only the story in pictures is "moved" from one system to another, whereas the story in language only "adds" the semiotic system of gestures. As we have seen, gestures in these two conditions were used differently in their interaction with speech, which supports the view discussed in Section 2.3.4 that language and gestures are two distinct but closely interacting semiotic systems (instead of constituting a single system). That is, spatio-motoric representations combine with linguistic representations for a better narrative expression suitable for each situation.

Chapter 5 Conclusions

This thesis centred around the *polysemiotic* and *multimodal* nature of human communication, which until recently has often been neglected in language related research. Considering the interplay of different semiotic systems and sensory modalities in human communicative behaviour, instead of examining these semiotic systems in isolation, more research needs to be conducted that takes into account how meaning is constructed in polysemiotic and multimodal interaction.

This study examined how narratives in different semiotic systems and sensory modalities are translated into embodied narratives, and investigated the influence of the source semiotic system on this process. The first research question at the start of the study asked what happens when the same story expressed in either language or pictures is translated into an embodied narrative. Since the two semiotic source systems have different potentials for expressing meaning, it was expected that some system-specific elements would transfer to the target narratives delivered in speech and co-speech gestures. Thus the two different "retellings" of the same story would lead to different multimodal polysemiotic narratives. This expectation was supported, and the source narratives translated into embodied narratives were different, for example, in length, in the use of gestures, and in terms of creativity.

The second research question asked whether hearing the story would result in more coherent embodied narratives in respect to organisation and the development of the plot. While the results do not indicate that the narratives translated from the speech narrative would result in more coherent embodied narratives than those translated from the picture narrative, they do, however, suggest some differences between the conditions in the use of devices that are used to create cohesion. Furthermore, more side comments (i.e. non-narrative clauses) were used in the speech narrative condition to support the development of the plot. Thus, it can be concluded that hearing a story does not necessarily result into a more coherent embodied narrative compared to when the story is seen in pictures.

The third research question asked if seeing the story would give rise to more perceptually detailed narratives, including iconic gestures and ideophones. Contrary to expectations, more iconic gestures were in fact found in the narratives translated from the speech narrative, but these were nearly entirely non-enacting in type. However, iconic first-person, enacting gestures were indeed significantly more frequent in the picture condition, and these enactments occurred often in connection with direct speech, which contributed to making the narratives of this group livelier. On the other hand, the proportion of ideophones did not significantly differ between the conditions, but ideophonic interjections that represented actions and reactions of the protagonists were only present

in the picture condition.

The reason for the differences in the use of iconic gestures was identified being that the narrators in the speech condition had an outside-the-story perspective in their narration, which required more systematic use of gesture space in order to represent developments of the events and track referents, whereas the narrators in the picture condition used an inside-the-story perspective, and were able to refer to things in relation to their own body. This condition-specific difference in perspective could be linked to the differences in the level of experientiality, which arose as the most dominant feature differencing the intersemiotic translations from one another. It was clear that the level of experientiality was lower in the speech narrative condition, reflected in fewer cases of first person perspective in both speech and gestural expression than in the narratives translated from the pictorial narrative. The embodied narratives in the picture condition often resulted in more personalised improvisations where the narrative "added" to the story. For example, narrators named the characters, motivated the introductory setting to the story, and provided events leading to action with background information. It was found that in many occasions the narratives translated from speech only merely provided information in a report style about the characters, the main events, and about the resolution, whereas in the picture narrative condition the retelling performance was more comparable to genuine storytelling. This difference can be traced back to the source narrative systems with language being stricter in terms of "accuracy", whereas the pictorial system affords a greater semiotic freedom of interpretation. In general, the story was not as well internalised in the speech condition, which was reflected in many cases of hesitation and uncertainty that affected how the story was narrated. Thus, although most of the specific hypotheses did not receive support in this study, the results nevertheless indicate that different source semiotic systems may indeed affect how polysemiotic narratives are constructed.

For future research, the non-enacting iconic gestures could be analysed further by dividing those that "depict" from those that simply "represent", allowing us to differentiate the gestures that reflect iconicity from those that are used to maintain coherence. Also, it would be worthwhile to extend this study by adding a third narrative condition: writing. This would allow us to see if writing would result in embodied narratives that are more similar to those translated from speech (as would be expected if both are subsystems of language), or to those translated from pictures (given that written stories are known to possess a high degree of experientiality, and are perceived visually). It could be seen whether the visually perceived subsystem of language would reflect greater narrative coherence, which in this study was expected from the subsystem of speech. This would further allow us to compare the role of the sensory modality better in the embodied narratives, and test the possible dominance of visual perception. Furthermore, to avoid having

language in both source narrative and target narrative, one way would be to present a simpler story in speech and in pictures, and ask the participants to renarrate the story in pantomime. Although that would not answer to questions about polysemiotic and multimodal communication per se, it could, however, reveal something about the role of sensory modality of hearing and vision that are two of the main sensory modalities we use in multimodal interaction.

References

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies 38*(1/4), 298-348.

Andrén, M. (2010). Children's gestures from 18 to 30 months. Ph.D. dissertation, Centre for Languages and Literature, Lund University. Retrieved from: http://www.lu.se/lup/publication/1700528.

Bacon, H. (2000). *Audiovisuaalisen kerronnan teoria.* Helsinki: SKS.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Berman, R., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study.* New Jersey: Lawrence Erlbaum Associates.

Chomsky, N. (1957). Syntactic structures. *Language 33*(3/1), 375-408.

Clark, H. H. (2004). Variations on a ranarian theme. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives*, 457-476. London: Lawrence Erlbaum Associates, Publishers.

Cohn, N. (2013). Visual narrative structure. *Cognitive Science 34,* 413-452.

DOI: 10.1111/cogs.12016

Coseriu, E. (1985). Linguistic competence: What is it really? *The modern Language Review 80(*4), XXV-XXXV. The Presidential Address of the Modern Humanities Research Association. DOI: 10.2307/3729050

Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. (2013). Gestural viewpoint signals referent accessibility. *Discourse Processes 50,* 431-456.

Dingemanse, M. (2011). The meaning and use of ideophones in Siwu. PhD dissertation, Nijmegen: Radboud University. Available from: http://thesis.ideophone.org/.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass 6*(10), 654-672.

Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics 3*(1): 4, 1-30.

DOI: https://doi.org/10.5334/gjgl.444

Donald, M. (1991). *Origins of the modern mind – Three stages in the evolution of culture and cognition*. London: Harvard University Press.

Efron, D. (1972). *Gesture, race and culture*. Preface by Paul Ekman. [Reissue of Efron 1941]. The Hague: Mouton and Co.

Fludernik, M. (1996). *Towards a 'natural' narratology.* New York: Routledge.

Gal, S. (1978). Peasant men can't get wives: Language change and sex roles in a bilingual community. *Language in Society 7*(1), 1-16.

Gibson, J. J. (1979). *Ecological approach to visual perception*. Boston: Houghton Mifflin.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science 12*(6)*,* 516-522.

Goodwin, C. (2003). The body in action. In J. Coupland and R. Gwyn (Eds.), *Discourse, the body and identity,* 19-42. Basingstoke: Palgrave Macmillan.

Graziano, M. & Gullberg, M. (2018). When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. *Frontiers in Psychology 9,* 1-17. DOI: 10.3389/fpsyg.2018.00879

Green, J. (2014). *Drawn from the ground*: *Sound, sign and inscription in Central Australian sand stories.* Cambridge: Cambridge University Press.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning 56*(1), 155–196.

Haddington, P., & Kääntä, L. (Eds.). (2011). *Kieli, keho ja vuorovaikutus: Multimodaalinen näkökulma sosiaaliseen toimintaan.* Helsinki: Suomalaisen Kirjallisuuden Seura. Suomalaisen Kirjallisuuden Seuran Toimituksia, Tiede, 1337.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: a comparison of English, French, German, and Mandarin Chinese. *Child Language 26*, 419-452.

Hockett, C. F. (1960). The origin of speech. *Scientific American 203,* 88-111.

Husserl, E. [1936] (1970). *The crisis of European sciences and transcendental phenomenology*. Evanston, IL: Northwestern University Press.

Hühn, P., Pier, J., Schmid, W., & Schönert, J. (Eds.) (2009). *Handbook of narratology*. Berlin: Walter de Gruyter GmbH & Co.

Jakobson, R. (1959). *On translation*. Brower, Reuben Arthur (Eds.). Harvard University Press. DOI: https://doi.org/10.4159/harvard.9780674731615.c18

Jakobson, R. (1965). Quest for the essence of language. *Diogenes 13*, 21-38.

Jääskeläinen, A. (2009). Mitä eletutkimuksessa tapahtuu? *Virittäjä 4*, 633-638.

Jääskeläinen, A. (2013). Todisteena äänen kuva: Suomen kielen imitatiivikonstruktiot. Helsinki: Unigrafia Oy.

Kajannes, K., & Kirstinä, L. (2000). *Kirjallisuus, kieli ja kognitio: Kognitiivisesta kirjallisuuden ja kielentutkimuksesta.* Helsinki: Helsinki University Press.

Kallio, A. (2010). Kertojat, kuulijat, katseet – Keskustelunanalyyttinen tutkimus vuorovaikutuksen multimodaalisuudesta kerrontasekvensseissä. Tampere: Tampereen yliopisto. Retrieved from: http://urn.fi/urn:nbn:fi:uta-1-20641

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35, 379-415. DOI: https://doi.org/10.1515/ling.1997.35.2.379

Kita, S. (2009) Cross-cultural variation of speech-accompanying gesture: A review. *Language and cognitive processes 24*(2), 145-167.

Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language 48,* 16-32.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication.* London: Routledge.

Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular*. Philadelphia: University of Pennsylvania Press.

Lakoff G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.

Leskinen, J. (2001). *Itämerensuomalaista ekspressiivisanaston tutkimusta*. Jyväskylä: Jyväskylän yliopisto.

Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes 15*, 277-301.

Mandler, J. M. (1984). Stories, scripts, and scenes: Aspects of schema theory. London: Lawrence Erlbaum Associates, Publishers.

Martinet, A. (1984). Double articulation as a criterion of linguisticity. *Language Sciences 6*(1), 31-38.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.

McNeill. D. (1992). *Hand and Mind: What gestures reveal about the thought.* London: The University of Chicago Press.

Menary, R. (2008). Embodied narratives. *Journal of Consciousness Studies 15*(6), 63–84.

Merleau-Ponty, M. (2002). *Phenomenology of perception*. (Colin Smith, Trans.). London: Routledge.

Mushin, I. (2001). *Evidentiality and epistemological stance: Narrative retelling.* Amsterdam: John Benjamin's Publishing Company.

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In *Body – Language – Communication: An international handbook on multimodality in human interaction 2*, 1687-1702. Berlin: De Gruyter Mouton.

Ortactepe, D. (2012). *The development of conceptual socialization in international students: A language socialization perspective on conceptual fluency and social identity.* Newcastle: Cambridge Scholars Publishing.

Payrató, L. & Teßendorf, S. (2014). Pragmatic gestures. In *Body – Language – Communication: An international handbook on multimodality in human interaction 2*, 1531-1537. Berlin: De Gruyter Mouton.

Piaget, J. (1962). *Play, dreams, and imitation in childhood*. New York: Norton.

Prince, G. (2008). Narrativehood, narrativeness, narrativity, narratability. In J. Pier and J. Á. García Landa (Eds.)*, Theorizing narrativity 19-27*. Berlin: de Gruyter.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – a window into perception, thought and language. *Journal of Consciousness Studies 8*(12), 3-34.

Ranta, M. (2011). Stories in pictures (and non-pictorial objects) – A narratological and cognitive psychological approach. *Contemporary Aesthetics 9*. Arnold Berleant.

Ranta, M. (2013). (Re-)Creating order: Narrativity and implied world views in pictures. *Storyworlds: A Journal of Narrative Studies 5,* 1-30.

Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In Feldman, R. S., & Rimé, B. (Eds.), *Fundamentals of Nonverbal Behaviour*, 239-281. Cambridge: Cambridge University Press.

SALT (2015). [Frog, where are you?: English transcript] *Systematic Analysis of Language Transcripts*. Retrieved from: http://www.saltsoftware.com/coursefiles/1202/FWAY_ English.pdf

Seppä, A. (2012). *Kuvien tulkinta: Menetelmäopas kuvataiteen ja visuaalisen kulttuurin tulkitsijalle.* Tampere: Gaudeamus.

So, W.C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science 33*(1), 115-125. doi:10.1111/j.1551-6709.2008.01006.x.

Sokolowski, R. (2000). *Introduction to phenomenology*. Cambridge: Cambridge University Press.

Sanford, A. J., & Emmott, C. (2012). *Mind, brain and narrative.* Cambridge University Press.

Sonesson, G. (1997). Mute narratives: New issues in the study of pictorial texts, Lund, May 1995. In Lagerroth, U-B, Lund, H., & Hedling, E. (Eds.), *Interart Poetics. Acts of the congress "Interart Studies: New Perspectives",* 243-252. Amsterdam: Rodopi.

Sonesson, G. (2010). From mimicry to mime by way of mimesis: Reflections on a general theory of iconicity. *Sign Systems Studies 38*(1/4), 18-66.

Sonesson, G. (2011). The Mind in the picture and the picture in the mind: A phenomenological approach to cognitive semiotics. In *Lexia. Rivista di semiotica 07/08*, 167-182.

Sonesson, S. (2014). Translation and other acts of meaning: In between cognitive semiotics and semiotics of culture. In *Cognitive Semiotics 7(*2), 249-280.

Sonesson, G. (2015a). From remembering to memory by way of culture. A study in cognitive semiotics. *Southern Journal of Semiotics 5*(1). *Special issue: Memory as a representational phenomenon*, 25-52.

Sonesson, G. (2015b). Semiotics of photography: The state of the art. In P. P. Trifonas (Eds.), *International Handbook of Semiotics,* 417-483. DOI: 10.1007/978-94-017-9404-6_19

Sonesson, G., & Lenninger, S. (2015). The psychological development of semiotic competence: From the window to the movie by way of the mirror. *Cognitive Development 36*: 191-201.

Streeck, J., Goodwin, C., & LeBaron, C. (2011). *Embodied interaction – Language and body in the material world*. New York: Cambridge University Press.

Tomlin, R. S., & Myachykov, A. (2015). Attention and salience. In Dabrowska, E. & Divjak, D. (Eds.), *Handbook of Cognitive Linguistics, 31-52.* Berlin: De Gruyter Mouton.

Trask, R. L. (2004). *Language: the basics (2nd ed.).* Oxon: Routledge.

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language 50*(4), 395-407. DOI: 10.1016/j.jml.2004.01.002

Wagner Cook, S. (2014). Gesture and working memory. In *Body – Language – Communication: An international handbook on multimodality in human interaction 2*, 1936-1940. Berlin: De Gruyter Mouton.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of LREC, Fifth International Conference on Language Resources and Evaluation*. Software available at http://www.lat-mpi.eu/tools/elan/

Yiheng Zhao, H. (2015). Editor's note. *Signs & Media SCU 11*, 2-3.

Zahavi, D. (2001). Beyond empathy. Phenomenological approach to intersubjectivity. *Journal of Consciousness Studies 8*(5-7), 151-167.

Zahavi, D. (2003). *Husserl's phenomenology.* Stanford, CA: Stanford University Press.

Zlatev, J. (2005). What's in a schema? Bodily mimesis and the grounding of language. In Beate Hampe (Eds.), *From perception to meaning: Image schemas in cognitive linguistics,* 313-342. Berlin: Mouton de Gruyter.

Zlatev, J. (2007). Embodiment, language and mimesis. In T. Ziemke, J. Zlatev, & R. Franck (Eds.), *Body, language, mind 1: Embodiment*, 297-337. Mouton de Gruyter.

Zlatev, J. (2008a). The dependence of language on consciousness. *Journal of Consciousness Studies 15*(6), 34-62.

Zlatev, J. (2008b). The co-evolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. P. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity,* 215-244. Amsterdam: Benjamins

Zlatev, J. (2009). The Semiotic hierarchy: Life, consciousness, signs and language. *Cognitive Semiotics 4*, 169-200.

Zlatev, J. (2011). From cognitive to integral linguistics and back again. *Intellectica* 56(2), 125-147.

Zlatev, J. (2012). Cognitive semiotics: An emerging field for the transdisciplinary study of meaning. *Public Journal of Semiotics 4(1)*, 2-24.

Zlatev, J. (2014). Image schemas, mimetic schemas, and children's gestures. *Cognitive Semiotics 7*(1), 3-29.

Zlatev, J. (2015a) Cognitive semiotics. In Trifonas, P. (Eds.), *International handbook of semiotics*, 1043-1067. Dordrecht, Netherlands: Springer.

Zlatev, J. (2015b). The emergence of gestures. In B. MacWhinney & W. O'Grady (Eds.) *The Handbook of language emergence*. Malden, MA.: Wiley-Blackwell Publishing.

Zlatev, J., & Blomberg, J. (2016). Embodied intersubjectivity, sedimentation and non-actual motion expressions. *Nordic Journal of Linguistics* 39(2), 185-208.

Zlatev, J., Sonesson, G., & Konderak, P. (2016). *Meaning, mind and communication: Explorations in cognitive semiotics.* Frankfurt am Main: Peter Lang GmbH.

Appendix A Frog, where are you? (adapted from Mayer, 1969)











APPENDIX B FROG, WHERE ARE YOU? TRANSCRIPT (ENGLISH)

(Adapted from Systematic Analysis of Language Transcripts (SALT), 2015)

Page	Script
1	There once was a boy who had a dog and a pet frog. He kept the frog in a large jar in his bedroom.
2	One night while he and his dog were sleeping, the frog climbed out of the jar. He jumped out of an open window.
3	When the boy and the dog woke up the next morning, they saw that the jar was empty.
4	The boy looked everywhere for the frog. The dog looked for the frog too. When the dog tried to look in the jar, he got his head stuck.
5	The boy called out the open window, “Frog, where are you?” The dog leaned out the window with the jar still stuck on his head.
6	The jar was so heavy that the dog fell out of the window headfirst!
7	The boy picked up the dog to make sure he was ok. The dog wasn’t hurt but the jar was smashed.
8	The boy and the dog looked outside for the frog. The boy called for the frog.
9	He called down a hole in the ground while the dog barked at some bees in a beehive.
10	A gopher popped out of the hole and bit the boy right on his nose. Meanwhile, the dog was still bothering the bees, jumping up on the tree and barking at them.
11	The beehive fell down and all of the bees flew out. The bees were angry at the dog for ruining their home. The boy wasn’t paying any attention to the dog. He had noticed a large hole in a tree. So he climbed up the tree and called down the hole.
12	All of a sudden an owl swooped out of the hole and knocked the boy to the ground. The dog ran past the boy as fast as he could because the bees were chasing him.
13	The owl chased the boy all the way to a large rock.
14	The boy climbed up on the rock and called again for his frog. He held onto some branches so he wouldn’t fall.

15	But the branches weren't really branches! They were deer antlers. The deer picked up the boy on his head.
16	The deer started running with the boy still on his head. The dog ran along too. They were getting close to a cliff.
17	The deer stopped suddenly and the boy and the dog fell over the edge of the cliff.
18	There was a pond below the cliff. They landed with a splash right on top of one another.
19	They heard a familiar sound.
20	They crept towards a big log and the boy told the dog to be very quiet.
21	Together they peeked over the log.
22	There they found the boy's pet frog. He had a mother frog with him.
23	They had some baby frogs and one of them jumped toward the boy.
24	The baby frog liked the boy and the boy and the dog were happy to have a new pet frog to take home. As they walked away the boy waved and said "goodbye" to his old frog and his family.

APPENDIX C FROG, WHERE ARE YOU? TRANSCRIPT (FINNISH)

Page	Script
1	Olipa kerran poika, jolla oli koira ja lemmikkisammakko. Hän piti sammakkoa isossa purkissa makuuhuoneessaan.
2	Eräänä yönä kun poika ja koira nukkuivat, sammakko kiipesi pois purkista ja hyppäsi ulos avoimesta ikkunasta.
3	Seuraavana aamuna kun poika ja koira heräsivät, he näkivät, että purkki oli tyhjä.
4	Poika ja koira etsivät sammakkoa kaikkialta. Kun koira yritti katsoa purkista, jäi sen pää purkkiin jumiin.
5	Poika huusi avoimesta ikkunasta, ”Sammakko, missä olet?” koiran nojatessa ikkunasta ulos purkki vieläkin päässään kiinni.
6	Purkki oli niin painava, että koira tippui pää edellä ikkunasta ulos.
7	Poika nosti koiran syliinsä tarkistaakseen, että se oli kunnossa. Koiralle ei ollut käynyt kuinkaan, mutta purkki oli säpäleinä.
8	Poika ja koira etsivät sammakkoa ulkona. Poika huusi taas sammakkoa.
9	Hän huusi maassa olevaan kuoppaan samalla kun koira haukkui mehiläisille mehiläispesässä.
10	Taskurotta tuli esiin kuopasta ja puri poikaa suoraan nenään. Samaan aikaan koira häiritsi vieläkin mehiläisiä, hyppi puuta vasten ja haukkui niille.
11	Mehiläispesä tippui maahan, ja kaikki mehiläiset lensivät ulos pesästä. Mehiläiset olivat vihaisia koiralle kotinsa rikkomisesta. Poika ei kiinnittänyt mitään huomiota koiraan. Sen sijaan hän oli huomannut ison reiän puussa. Niinpä hän kiipesi puuhun ja huusi reikään.
12	Yhtäkkiä pöllö syöksähti reiästä ja tuuppasi pojan maahan. Koira juoksi pojan ohi niin kovaa kuin jaloistaan pääsi, koska mehiläiset jahtasivat sitä.
13	Pöllö jahtasi poikaa aina suurelle kivelle saakka.
14	Poika kiipesi ylös kivelle, ja huusi taas sammakkoaan. Hän piti tukea oksista jottei putoaisi.
15	Mutta oksat eivät oikeasti olleetkaan oksia! Ne olivatkin peuran sarvet. Peura nosti pojan sarvilleen.
16	ja lähti juoksemaan poika sarvissaan. Koira juoksi mukana myös. He lähestyivät jyrkännettä.

17	Peura pysähtyi äkkiä, ja poika ja koira tippuivat jyrkänteen laidan yli.
18	Jyrkänteen alapuolella oli lammikko. Poika ja koira läiskähtivät toinen toisensa päälle lammikkoon.
19	Nyt he kuulivat jostain tutun äänen.
20	He hiipivät suuren kaatuneen puunrungon luo ja poika käski koiraa olemaan hyvin hiljaa.
21	He nousivat yhdessä kurkistamaan puunrungon yli.
22	Sieltä he löysivät pojan lemmikkisammakon. Sillä oli toinen sammakko mukana.
23	Niillä oli myös poikasia, ja yksi niistä hyppäsi poikaa kohti.
24	Vauvasammakko tykkäsi pojasta, ja poika ja koira olivat onnellisia saadessaan viedä uuden lemmikkisammakon kotiin. Kävellessään pois poika vilkutti ja sanoi ”näkemiin” vanhalle sammakolleen ja sen perheelle.

APPENDIX D FEEDBACK FORM

(Translated from Finnish)

1. Gender: _____
2. Age: _____
3. Are you left or right handed? _____
4. Languages spoken rating from 1 (beginner) to 5 (advanced): _____

   _____

5. In your everyday life do you regularly use some other language in addition to Finnish (home environment/leisure/work/education)? If yes, what language and in which situations?

   _____

   _____

6. Have you lived abroad longer periods than 6 months? If yes, where, how long, and at what age?

   _____

   _____

7. Did you consider the experiment challenging? If yes, why?

   _____

   _____

8. Do you think the task could have been easier or alternatively more challenging if instead of hearing the story you would have seen the story in pictures (without language) / seeing the story you would have heard the story spoken on tape? If yes, please indicate the reason for your answer. _____

   _____

9. Do you consider yourself visual or auditory learner (i.e. you learn better via visual perception, or auditory perception)?

   _____

   _____

Thank you for your participation!