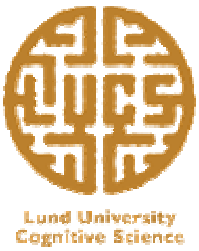

Talmy's Manner in Event Perception

An eye-tracking approach to linguistic relativity

Master's Thesis 20p / 30ECTS
Student: Richard Andersson
Supervisor: Kenneth Holmqvist
May, 2005

ABSTRACT
This study uses eye-tracking equipment to search for effects of linguistic relativity. The study tries to escape the traditional battle between anti-relativistic modularism and pro-relativistic connectionism by testing effects of "distributed" linguistic relativity. It finds that Talmy's manner element has a perception-attracting quality which guides attention in unequal amounts due the different manner density in S- and V-languages. The attracting area is the manner's active zone in the figure. Furthermore, this change in attention makes way for differences in event memory. This was not *directly* proven in this study, but is suggested by indirect memory results.

ACKNOWLEDGEMENTS
I would like to thank the following people for their support during the work of this Master's thesis (in no particular order):

Kenneth Holmqvist – for supervising and putting in long hours in the lab during my testing.

LUCS eye-tracking group, spring 2005: Philip Diderichsen, Andreas Dren, Johanna Fröjmark, Erik Johannesson, Marcus Nyström and Stefan Visuri – for invaluable critique and feedback.

Simon Retailleau & Valérie Laazizi – For translations and unorthodox French lessons.

Joost van de Weijer – For help with the statistics.

Sven Svenström – for ideas during the hypothesis formulation. Sorry it never turned into a frog-story study.

Jordan Zlatev – for inspiration and getting me hooked on semantics and relativity.

Lhinn Nilsson-Nyman, Helena Lidén & Peter Magnusson at Delphi XY3 – for modelling.

INDEX

1. INTRODUCTION

Linguistic relativity is a controversial subject and people tend to have strong opinions about it. Perhaps the idea that we are unconsciously determined by language is the part that most people object to, because nobody want to be stripped of their free will. However, the power of linguistic relativity is not very strong and it does not, as far as we know, force people to think in terms they do not want to use. Still, it is an exciting area which have yielded several interesting results the past decade.

This study searches for effects of linguistic relativity based in how languages encode the manner in which a movement is carried out. More demanding forms of expressing the manner means an increased chance of leaving such expressions out of the utterance. This has a consequence to the hearers of the utterances who never hear the manner expressions and this may affect their perception of the event. This study makes two different experiments: one scene perception experiment and later a memory recall experiment. The experiments are conducted using eye-tracking equipment.

2. LINGUISTIC RELATIVITY

2.1 The old relativity

The notion of linguistic relativity is about how the language we use influence how we dissect reality and ultimately how we think. This idea is by no means a new one, as writings of the German linguist Wilhelm von Humboldt in the late 1700s reveal (Allwood 1983). However, the principle of linguistic relativity reached new heights of controversy and fame through the work of Benjamin Lee Whorf who presented the radical new claim that Hopi indians, through the structure of their language, perceived reality radically different from speakers of "Standard Average European" (SAE). To make an extreme simplification and abbreviation of Whorf's claims: speakers of SAE experience "things" whereas speakers of Hopi experience "events" as a consequence of the linguistic patterns Whorf called "habitual thought" (Whorf 1956: 147).

Whorf's ideas were greatly debated and in the end more or less discarded as unsubstantiated speculations. This, however, did not prevent this entire notion of linguistic relativity to become known as the Sapir-Whorf hypothesis, the Whorf hypothesis and the Whorfian hypothesis. In general, a lot of confusion exists about what Whorf did or did not claim. The term "hypothesis" is most likely not an invention of Whorf, but of others, supposedly by Eric Lenneberg (according to Brown (1976), in Alford (1980)). Also the distinction between a "weak" and a "strong" version, is something not originally mentioned by Whorf (Alford 1980). Linguistic relativity was the idea that the linguisitic differences between two language structures would also be correlated with non-linguistic differences, implying that different speakers of a language would experience reality differently from speakers of another language. The idea of linguistic determinism was that these differences were also strongly or completely determined by the particular language (Alford 1980).

2.2 The neo-Whorfianism[1]

Today the terms seem to have merged together under the common term "linguistic relativity" but the aim of the research is that of determinism, i.e. to give evidence that the non-linguistic cognitive effects are influenced by language (Lucy 1997). However, the requirement for the effect to be strongly or completely determined by language has been abandoned. The current field seems to suggest that that there may be a whole continuum of effects with varying degrees of influence ranging from weakly influencing to strongly determining. The new

[1] A term taken from Levinson (2003: pp 301-307) to denote the new rise of interest in linguistic relativity.

relativity studies have made an effort to more precisely specify what variable in the language actually influences what variable in cognition. Slobin (2003) presented his "thinking for speaking"-paradigm, which basically means that relative effects of language stem from the fact that different parts of the particular language do not function in the same way when using them to encode reality. In other words, just looking at something does not give a different experience if you have a different language. However, if you actively use the language in the task, such as when telling a story, then effects of linguistic relativity may be elicited. Slobin focused on the production side of language by making subjects tell a narrative based on pictures, and then focus on if this structurally differed between languages. The vital part of this was that the attention was language-mediated, because different languages express different things with different ease, therefore letting the language guide attention (for an example of language-mediated effect only, see Feist & Gentner 2001). This is shown in "frog story studies", which using a children's picture book to elicit the stories from the subjects. Research using these frog stories showed that speakers of different languages mention different semantic elements of the stories. Some speakers mention the manner in which a character moves, whereas other speakers merely mention the path along which the character travels (for example, Slobin in Strömqvist & Verhoeven 2004).

However, the interesting and perhaps controversial part is non-linguistic cognitive consequences from the linguistic processing – if language processing has some effect on the mind that carries over into other functions of the mind. As Bloom and Keil (2001) put it:

> *Before getting to all this, however, we should point out that the issue here is not about whether language can have an effect on thought. Of course it can. (If it couldn't, why would you be reading this?). Nobody doubts that language can inform, convince, persuade, soothe, dismay, encourage, and so on. This is what language is for.*
> …
> *The debate, as we see it, is not whether language shapes thought – it is whether language shapes thought in some way other than through the semantic information that it conveys.*

Several ways of exploring the cognitive consequences of language are currently being researched. Some of the more well-known are how language influences our spatial frames of reference (Levinson 2003), imagery (Slobin, 2003) and memory (Oh, 2003. In Slobin: In press.).

3. TALMY'S COGNITIVE SEMANTICS

The linguist Leonard Talmy's semantic elements have been greatly used by, for example, Slobin (2004), Levinson (2003) and Kita (2003) in cross-linguistic studies examining how space is treated by language and thought. These semantic elements are part of Talmy's attempt to create a cognitive semantics which tried to explain semantics with the mind as a point of departure. Of these elements, some are more widely used and accepted than other. The relevant elements which are used in this study are explained below. These elements are purely semantical units used to denote particular meanings in especially motion events. These elements are often tied to specific words, but can equally well be carried at sentence level, by expressions or at other syntactic levels. The relevant semantic elements for this study are figure, ground, motion, path and manner. These we be elaborated on below.

3.1 Figure and Ground

The *figure* is the most relevant object in an event and it functions as the "main character" in an event. In the following example:

(1) The car raced past the police patrol.

The car is the "main character" and the highlighted object. It is used as a subject in this sentence, it is the agent and it is uttered in active voice, which serve to highligh it as the important part. The car carries the semantic element known as the figure.

The police patrol, however, is not a figure, but merely uttered as a reference object to the movement of the car. This function of being the reference object to another object and its movement is called being the *ground*. Not only is it possible to pick out a figure and a ground in a sentence, but also in a picture and in an image schema (see figure 1). The ground is a basically an element used to describe that a particular part carries the meaning of a reference object.

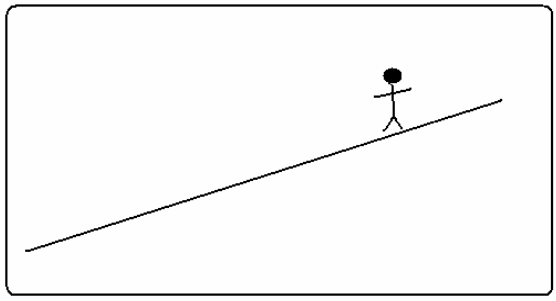

Figure 1: In this picture, the stickman is identified as the main character and important part, whereas the hill is just a stage for the main character to place him in a setting. In other words, the stickman is the *figure* and the hill is the *ground*.

The figure and the ground are closely tied to each other since it is hard for either to exist without the other in a motion event. This is because movement is per definition a change in position between two objects, and therefore we cannot preceive the movement unless it is in relation to something else. In example (1), the use of the patrol car was necessary to make the fact that the car raced past something apparent.

A ground is not always overtly present, for example:

(2) John showered

However, the ground is an element of meaning, and therefore it does not have an overt 1-to-1 matching with words. In example (2), it can be argued that the ground is present in the verb in the form of a representational image of a shower. John is thus the figure located in the ground known as the shower.

For motion events, which are the events studied in this thesis, omitted ground produces unnatural-sounding sentences unless it is evident from context what the ground is. For example:

(3) *John ran past

It is of course possible to think about possible situations where this sentence would be valid. Example (3) could be possible in the following utterance (omitted ground in parentheses):

(4) I think I just saw John running past (me/here).

Considering another sentence:

(5) ?John raced

In example (5), there is a semantic shift when not using the ground. It is no longer a question of 'John racing past something', but rather 'John is currently racing (running in a limited area, such as in a stadium).

It is not at all clear-cut what is the ground and what is the figure in all possible utterances. Talmy (2003a: 315) gives a good summary of what differences to search for when trying to identify the figure and the ground:

	Figure	**Ground**
Definitional characteristics	Has unknown spatial (or temporal) properties to be determined	Acts as a reference entity, having known properties that can characterize the Figure's unknowns.
Associated characteristics	More movable	More permanently located
	Smaller	Larger
	Geometrically simpler (often pointlike) in its treatment	Geometrically more complex in its treatment
	More recently on the scene/in awareness	More familiar/expected
	Of greater concern/relevance	Of lesser concern/relevance
	Less immediately perceivable	More immediately perceivable
	More salient, once perceived	More backgrounded, once Figure is perceived
	More dependent	More independent.

Table 1. Differences between Figure and Ground. From Talmy (2003). *Toward a Cognitive Semantics – Vol I.* pp 315-316.

3.2 Motion

Motion simply means that the figure is in motion. The motion element does not however specify in what direction the motion is going or otherwise specify the nature of the it. An example of a simple expression of motion would be the verb "moving" as in:

(6) John is moving

Here the expression denotes motion, but it does not specify in what way the motion is carried out, nor in what direction the motion is going. This is specified by other components described in the following sections.

3.3 Path

The semantic element of *path* is an invisible trajectory or line along which the figure moves. If someone "enters a room", it would indicate a path which goes from the outside of the room, through the opening (typically the door) and into the room. Many verbs and satellites carry this information of path, for example the words "ascend" and "exit", which indicate a path "upwards" and a path "out of", respectively (Talmy 2003a: 99-116).

3.4 Manner

This semantic element signifies a "co-event" with the original motion event. The component describes the nature of the particular motion and how the motion is carried out. The example, the verb "rolled", not only signifies a motion of something, but also specifies the particular rotating way the object is moving (Talmy 2003b: 21, 27-33).

4. TALMY'S SEMANTIC ELEMENTS IN LANGUAGE TYPOLOGY

Talmy's elements have been found to differ in use in certain languages. Some languages *conflate* different elements differently, which means the languages "package" the semantic elements differently in words and expressions (Talmy 2003b: 28). For example, the highlighted verbs of the following sentences differ in conflation pattern:

(7) John **travels** to London

(8) Jacques **sort** de l'immeuble. (Jacques leaves the building)

The verb "travel" encompasses only the element of motion. The French verb "sort", however, encompasses both motion and an element of path. Not only does it mean that Jacques is moving, but also in what direction. In this case, it is the building, but a more general way of expressing it would be to say that the path of "sort" is simply out of the "container" which currently holds Jacques (Talmy 2003a: 196, 217).

Slobin (In press) makes frequent use of the distinction between S-languages (satellite-framed languages) and V-languages (verb-framed languages) and how they habitually treat the semantic elements. S-languages frequently uses the main verb to carry the semantic elements of motion and manner and an extra particle (called "satellite" by Talmy) encodes the path element. For example:

(9) Susan ran from the fire
("ran" = MOTION + MANNER, "from" = PATH)

The difference with V-languages is that they frequently encode the path element in the main verb and produce the manner element in a separate clause. For example:

(10) La fille descend la pente en glissant
[the girl descends the hill sliding]
"The girl slides down the hill"
("descend" = MOTION + PATH, "en glissant" = MANNER)

The implication of V-languages tending to produce the manner element in a separate clause, is that this part is often omitted in production to reduce cognitive costs, since an extra clause requires more "processing power" than a satellite or a main verb which get the manner conflated for free (Slobin 2003: 3, *ibid* 2004: 226).

Note however, that this distinction is by no means black and white, and the dichotomy is not clear-cut. Slobin rather speaks of a "cline of manner salience", where several factors work together to make manner more or less accessible for different languages (Slobin 2004: 250). Such factors may be linguistic practice (possibly based in the fact that the construction is cognitively demanding) and that a V-language may be constrained from making the particular construction (Papafragou *et al* 2002: 195-196). This constraint is known as the "boundary-crossing constraint" and means that V-languages are, to a greater extent, not allowed to use manner verbs if the path of the figure crosses a border. A border would for

example be going inside a house or diving into a pool, i.e. crossing the border of the ground. Going back to accessability: from a production perspective, accessibility translates into cognitive costs where a high accessibility means lower cognitive costs.

Of interest for this study is that Swedish (a S-language) has higher manner salience than French (a V-language), and generally encodes manner by habit. French, on the other hand, does not have manner equally accessible, and is therefore more likely to not encode the manner element in a perceived event. The implications for the hearers of Swedish and French is that they receive an unequal amount of descriptions containing the manner element. If this manner element affect the perception, then different attention-guiding patterns exist within the two languages and quite possibly within each of the larger S- and V-groups of languages.

5. EYE MOVEMENTS AND ATTENTION

5.1 Eye-tracking

This study uses eye-tracking equipment to monitor the subjects' eyes and their distal counterpart, i.e. what the subjects are looking at. Attention is the primary variable investigated in this study and therefore it is crucial that the eye-movements correlate well with attention. Although there exists, to my knowledge, no method of *directly* measuring the actual attention. Suggestive albeit indirect results in favor of a link between eye-movements and attention are abundant in both older and newer research. For an overview of the most relevant research see Rayner (1998). For a link between eye-movements and attention from a neurological perspective see Corbetta *et al* (1998). An important part of this experiment is how attention shifts according to the spoken linguistic stimuli, and both old and new research show that subjects fixate the visual referents of a spoken word (Cooper 1974. Tanenhaus, *et al* 2000). Not only are the visual counterparts of a word fixated, but also semantically related objects are perceived (Huettig & Altmann 2005), which is important to this study since it investigates semantic elements.

Not only is it crucial to know that the attention and the eye-movements correlate, but also that the eye-movements reflect the task at hand and are ecologically valid. In this regard, instructions are very important to be sure the results exist outside the laboratory. The russian psychophysicist Yarbus showed this using eye-tracking and varying the instructions (Yarbus 1967). An illustration (figure 2) from Yarbus (1967: 174) shows how disparate results can be achieved by varying the instructions:

In picture 1, the subjects where told to look freely on the picture. In the following pictures the tasks where the following: 2) estimate family wealth, 3) estimate ages of the members, 4) guess activity of the family prior to the visit, 5) to remember the clothes, 6) to remember the positions of the people and 7) to estimate how long the visitor has been away from the family. This shows that it is important to closely consider the

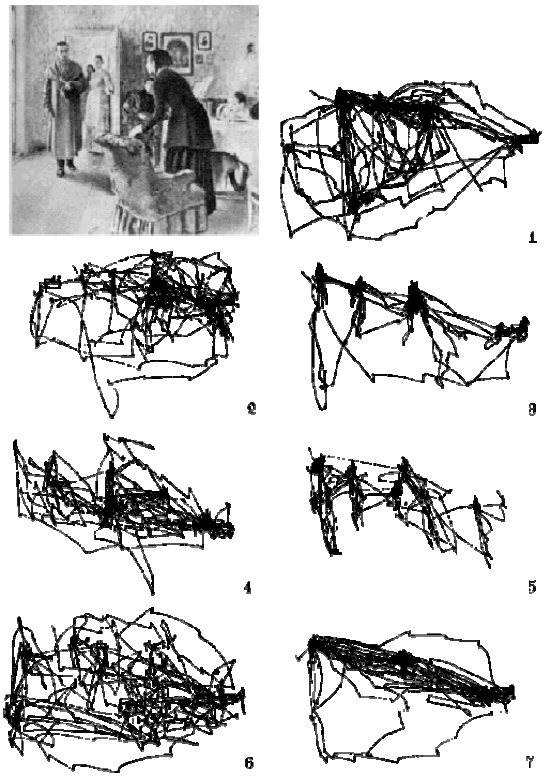

Figure 2: (Yarbus 1967: 174)

instructions given to the subjects to not influence them in favor (or exceedingly against) the tested hypothesis.

The validity of the instructions used in this study will be discussed at the end of the paper, in section 9.2.

5.2 Eye movements and memory

Not only are eye-movements highly correlated with attention, they are also well correlated with memory. In most cases, it is simply a prerequisite to have seen something first to be able to recall it later. Many studies from many different areas confirm this, for example from newspaper reading (Lundqvist & Holmqvist 2004), scene perception (Melcher & Kowler 2001) and picture processing (Nelson & Loftus 1980, in Henderson *et al* 2003: 726).

Memory tests are a useful approach to linguistic relativity because it is one of the non-linguistic areas of cognition which is fairly measureable and thus it is possible to detect cognitive effects influenced by language.

6. THE THESIS AND EXPERIMENTS – THE PROPOSED MECHANISM

Some studies investigating linguistic relativity have found manner effects on memory (Oh 2003, in Slobin: in press) and some have failed to find this at all (Papafragou 2002). Some argue that language and the rest of cognition are not that closely tied. Rather, the spoken description is just a limited package of our thoughts. Just because we express something a certain way does not mean we are limited to thinking about it or remembering it in that way (*ibid* : 212-216). This study does not aim to resolve this debate, but rather to look at linguistic relativity as a larger pattern of language and cognition use.

A way to work around the objection above is to look at the production-perception-recall process as distributed. The objection mentioned is that linguistic representations are paired with non-linguistic representations cancelling out any effects of relativity. However, this is not valid in a distributed language use. A speaker transmits the linguistic representation only and keeps the possible non-linguistic representation to himself/herself. However, the hearer only receives the linguistic representation and thus has to guide his/her attention accordingly. The attention in turn affects the memory since we, naturally, are better at remembering what we look (longer) at. This means that linguistic relativity may not only be a question of effects isolated in one human, but also at a higher level in a language community – a sort of "distributed linguistic relativity".

This thesis investigates the perception of motion events through the use of eye-tracking equipment. The point of departure is studies showing that speakers of verb-framed languages, such as French, are less likely to produce sentences which mention the manner in which a motion is carried out, compared to speakers of a satellite-framed language, such as Swedish. This, in turn, means that fellow speakers of the V-language hear reports of events which contain fewer descriptions of the manner. The question this thesis seeks to answer is whether these non-manner descriptions affect the perception of the hearers differently than descriptions containing manner expressions. If the notion of linguistic relativity is to be applied, these linguistic differences must also lead to cognitive differences. Therefore, this thesis examines whether these possible differences in perception may lead to differences in memory recall, since eye-tracking studies have shown that the longer you look at something, the better you are at remembering it (as explained in section 5.2).

This hypothesis can be illustrated in the following way:

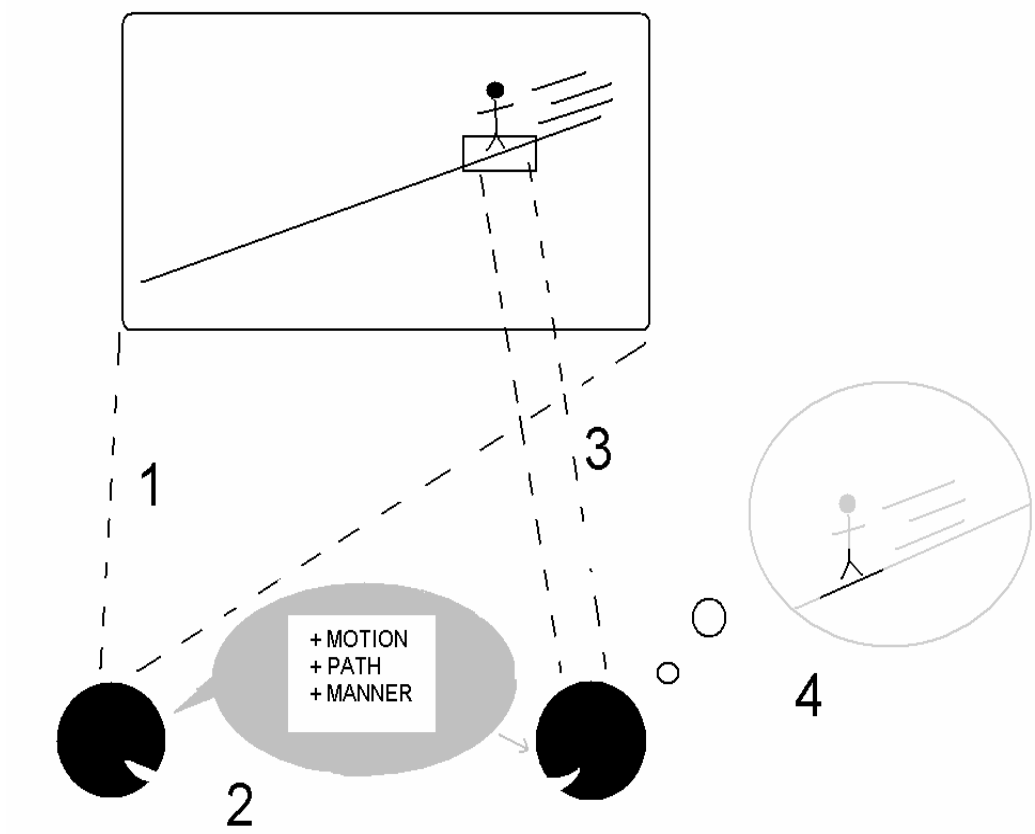

Figure 3: The proposed mechanism of linguistic relativity.

Figure 3 illustrates four different parts. First, a person perceives an event. Secondly, the person utters a description of it which may or may not contain the manner element. Thirdly, a hearer processes this utterance and looks at the same event. His/her perception is now guided differently based on if the manner element was part of the utterance or not (which it is in figure 3). The picture shows a manner utterance and as an effect the hearer focuses relatively more attention on the area of the figure responsible for the manner. Fourthly, when later remembering the event, the manner is relatively better recalled than for hearers who heard an utterance without the manner element.

The question is what specific differences in perception can be expected because of the existence or non-existence of a manner expression? How will a sentence using the description "ran down" compared to "went down" affect the attention and perception of the hearer? This thesis proposes that Talmy's semantic element of manner (and potentially every semantic element) channels the attention of the perceiver based on a "linguistic attention map"[2], which is different for every manner element, but has shared areas with all manner elements to a certain extent.

The "linguistic attention maps", which to some degree are idiosynchratic, are representations of an event which gives the hearer a map with information about the most salient parts of the event. In more concrete terms, a linguistic attention map of the manner element of "run" would for example attentionally highlight the parts handling the motion, i.e. the legs. The use of a verb without a manner element ("descend") would result in the

[2] This term is used throughout this section. It should be noted that the ideas here under section 6 are a proposal for a mechanism and not something proven or accepted, unless they are given proper reference to.

exclusion of the attention map of the manner element, and in turn less attention would be focused on the legs in this case. This study claims, and will examine, if the attentionally highlighted area by the manner element is the same area as the parts of the figure responsible for the motion. This area, the one responsible for the manner, is what Langacker would call the "active zone" of the manner element. An active zone is the particular part of a figure or ground involved in the described action (Langacker 1991: 189). For example:

(11) Simon kicked John.

In (11), the active zone of Simon would be the foot or leg, since it is used for the kick and it serves as a contact area between the two persons. The suggestion that the manner element maps on to the concept of active zone is in line with research showing that "language does not 'map' onto perceptual experience directly, but rather onto intermediary concepts" (Sigurd & Zlatev: In press). Also, this study uses the term "manner" in the original and general way Talmy suggested (Talmy 2003b: 21, 27-33). However, Dodge and Lakoff (In press) has divided up this term into four semantic strutural elements: Gait, Speed, Effort and BodyPart. This study does not seek to test which one of these elements correlate better with changes in attention. Furthermore, it does not seem possible to vary these elements in a systematic way. Speculatively, the element which seems closest to this study would be the BodyPart element. This will not be tested, however, as it seems extremely hard to vary by constructing sentences which use the element of Speed, but not that of BodyPart.

The manner element has its own linguistic attention map, which focuses the attention to the active zone. A linguistic attention map illustrated:

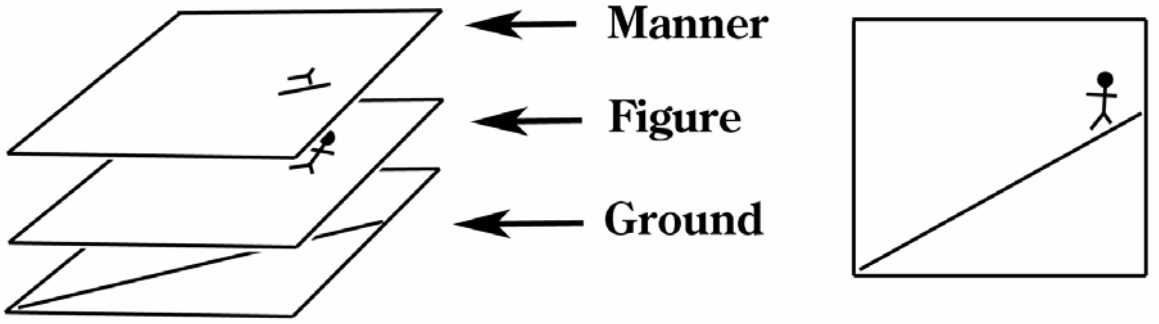

Figure 4: Linguistic attention maps

Figure 4 shows three different linguistic attention maps, one for every semantic element. The more areas that overlap in these different maps, the more attention these areas will attract. In picture 4, the attentional area of the manner element is added to the attention already attracted by the figure and the ground. These linguistic attention maps are then stacked to one final map which is a guide to the areas that will attract the most attention in a scene.

The reason for calling this linguistic attention distribution a "map" is that it seems simplistic just to say that the manner element focuses the attention on one part of an event. Rather, it could be so that a manner element such as the one in "He hurried into the cave", contains a particular distribution of attention. It is not unlikely that a perceiver may look at not only the legs, to see the increased motion of "hurry", but also the face of the person to confirm that he "hurries" in a psychological sense and not just moves fast because of his normal movement pace happen to be fast. Therefore, these linguistic attention maps show the complete attentional distribution of a semantic element across an entire event. However, this study will not focus on testing the whole attentional distribution of all semantic elements, but

focus on the attentional area of manner elements in motion events and, in particular, the area they are thought to have in common: the active zone of the manner element.

The task will be to see whether the attention of the perceiver shifts in the way previously described. The experiment will use manner elements which are thought to be located in a particular part of the figure (i.e. the legs) and not move about in the event. This is why a manner element such as the one in "roll" would not work, since it assumes the whole figure rolling and thus it would be impossible to determine whether a fixation is on the manner element or just on the figure.

To summarize the hypotheses of this thesis:

- H1: The semantic element of manner focuses attention to the active zones of the figure and the ground.
- H2: This attention results in a longer duration of total eye fixations in the active zones.
- H3: The longer durations of total eye fixations in the active zones will result in a better memory recall for details located in that area.

7.. METHOD

7.1 Eye-tracking terminology

The following definitions of the eye-tracking related terms were used: A fixation is any positions of the eye which give a centre of view in an area which is no greater than 5 mm in radius on the monitor used (~1.5º visual angle) and are retained in this area for at least 100 ms. An area of interest (AOI) is an area in which all fixations are counted and the specific area of the AOI is determined by the experiment leader with regards to where the theory predicts the measured effect will appear. A gaze is the sum of all fixations within an AOI which are carried out without leaving the AOI. Therefore, if the subject produces a fixation outside the AOI and then re-enters it, it will count as a second gaze. Furthermore, the experiment will make use of the term "fixation total" which is the combined time of all fixations in an AOI, even though they may belong to different gazes. This terminology is in agreement with the terminology of the field of eye-tracking (Jacob & Karn. 2002).

7.2 Equipment, design, setup and subjects

The eye-tracker was a head-mounted iView X made by SMI. This system also featured head-tracking to counter for any movements by the head and made it possible to produce a datafile with exact coordinates and times (as opposed to a video-file which must be analysed manually). The eye-tracking system measured the position of the eye every 20 milliseconds (i.e. 50 Hz). The eye-tracker was monitored and controlled from a computer in another room, where the eye-tracking specialist of this project made certain the camera obtained an unobstructed image of the eye throughout the experiment.

The stimuli was presented on a computer using E-prime, which is the program used to present the pictures and spoken sentences, as well as controlling when the eye-tracking computer should start and stop its recording. The computer was a 1.8 Gz P4 with 512 Mb memory. The computer's 19 inch monitor used a resolution of 1280x1024 pixels at 70 hz. All pictures were scaled to use the full size of the monitor. The subjects were positioned 50 cm away from the monitor on average.

A separate computer ran the recall test, which was also made with E-prime. This system was located in another room, approximately 20 meters from the eye-tracking room, so the subjects would be more likely to have cleared their working memory of any images as they started the recall test.

The datafiles produced by the eye-tracking system was analysed using iView Analysis and Excel. The statistical tests were performed using Excel and SPSS.

The subjects were 11 native French-speakers (exchange students) and 33 native Swedish-speakers. No control for gender or education level was made, which meant that the subjects were predominantly university level students. However, a fair balance of male and female subjects was obtained (27 male, 19 female). All French subjects except two were exchange students at Lund University. Both groups of subjects had by estimation a mean age of around 24 years, but both groups also had representatives from an older age group (40-45 years of age). The subjects knew that we filmed their eyes and some might have inferred from the calibration routine that we were able to see where they were looking. However, they were not told anything about the purpose of the experiment until after the tests. The data of one subject, who is not accounted for in the above subjects (11 + 33 subjects), was discarded when it was discovered that he had heard about the purpose of the experiment. All French subjects and 24 Swedish subjects performed the test after having participated in a production test before. The other 9 Swedish subjects were tested at a separate occasion where they did not receive the same narrative production test before the experiments. Instead, they conducted a test where they had to guide their attention in a controlled manner. It is likely that this influenced their eye movements to be not as free as the other subjects and their results are against the hypothesis of this paper to a larger degree. However, they are included in the data since their presence did not disrupt the significances of the statistical tests too much. It seems the added benefit of these extra subjects in the form of lowered variance was paired with their lower performance, thus not affecting the p-value of the statistical tests too much.

The subjects were tested in the Humanities Lab (Humanistlabbet) at Språk- och Litteraturcentrum (SOL-centrum) at Lund University. The subjects left electronic devices outside the test room to reduce any sources of disturbances to the equipment. The subjects were then positioned in front of the stimuli computer and received some quick information about the test. They were asked to listen to a couple of sentences immediately followed by a picture. Their task was to estimate how well the sentence and picture agreed. This procedure was repeated four times.

The subjects were then calibrated with the eye-tracking equipment and left to complete the experiment by themselves. All instructions for the experiment were inside the test, so this meant that Swedish subjects received Swedish instructions and French subjects received French instructions. This was deemed necessary so they would not be primed to any other language, e.g. English, which was the language used to communicate with the French subjects.

Immediately before the subjects started the two experiments, most of them first participated in an oral production experiment which was excluded from this thesis for difficulties with the analysis. However, the procedure will be accounted for here to give a full account of possible priming stimuli.

The subjects were shown four pictures from Mercer Myer's children's book *Frog, Where are you?*, which were picked out as they constituted a mini-story within the larger story of the book. They received instructions in their own language and controlled when to see the next picture of the story.

The subjects were first shown a quick slide-show of the pictures to give them a general idea of the story and the characters. Then they were given the instruction to tell a story based on these pictures starting with the first picture. The subjects were able to skip to the next picture by pressing spacebar at any time and were free to tell as much or as little as they wanted. After the fourth and last picture, the subjects were thanked and then received instructions for the next experiment which is the experiment described in this study.

7.3 Experiment I

7.3.1 Hypothesis

If the presence of a manner element attracts the attention to the active zone of the figure, then the following is expected:

- A longer duration of the average fixation total for the active zone for the subjects receiving sentences containing manner elements.

7.3.2 Stimuli

The pictures showing the motion events were 1280x1024 pixels and depicted the following things: Event 1 shows a young woman standing in the middle of a hill as if she is walking or sliding down the hill on her feet. Event 2 shows a young woman in profile as she is walking on a road. Event 3 shows a young woman running from a house. Event 4 shows a young man sneaking into a room. Since these are mere pictures, the fact that it is a *motion* event is not immediately obvious. Certain clues can be used to conclude that it is a motion event, namely, the speaker voice mentions the motion, the angle of the body (leaning forward suggests forward motion), and the position of the feet (weight distributed on one foot while the other is in the air, for example). It is possible that there are other clues to the motion, but these are mentioned as an example of the three most obvious ones. It should be mentioned that this ought not to affect the experiment even though it measures the fixation duration of the area around the feet. A person may look at the feet to check for motion and not for manner, but this is equal for both the non-manner and the manner-group. The manner group is expected to look *more* on the relevant area than the non-manner group. The event pictures along with the complete sentences will be presented together with the results in section 8.1. They can also be viewed in Appendix I.

The sentences were varied in terms of semantic structure to see if the observed effect would be applicable across a wide number of sentences. This semantic structure is accounted for in table 2:

Event number	Overt ground	Boundary-crossing	Direction of movement
1	Yes	No	Vertical
2	No	No	Horizontal
3	Yes	Yes	Horizontal
4	Yes	Yes	Horizontal

Table 2: semantic structure of the sentences

A clarification of table 2 is in order. Overt ground means that the ground is mentioned in the sentence and does not have be inferred from context or lexical meaning. Boundary-crossing refers to the constraint which applies to most V-languages (see section 4). Also, an event which had vertical movement was included to see if the results of this event would differ from the other in any way. It has been suggested that vertical motion is treated differently than horizontal motion and therefore (Zlatev & David 2004). If the effect is present in all events, it will also be present in many other sentences of the same semantic structure. This means the effect is not an isolated phenomenon, but something which is present in the language at large.

The sentences were recorded in 44 kHz, 128 kbit, and were spoken with no particular focus on any word. The French speaker had no particular dialect (i.e. spoke a standard variant), and the Swedish speaker spoke a Scanian dialect of Swedish.

7.3.3 Procedure

The subjects received either a manner version or a non-manner version of this experiment. The manner version had spoken sentences that mentioned a manner of the motion, whereas the non-manner version mentioned the motion in a non-manner expression.

The subjects were shown an instruction screen where they were told to rate the fit/validity between a spoken sentence and a picture. The exact instructions can be read in Appendix II. As soon as the subjects pressed spacebar to continue a small fixation cross appeared in the center of the screen, which the subjects had been told to look at. The purpose of the fixation cross is to let all subject have the same point of departure for their eyes. This fixation cross was located outside the target area of interest to eliminate the chance of "carry-over fixations" as the pictures of the events replaced the fixation cross on the screen. The fixation cross was presented during the time the spoken sentence was read and was followed by the appropriate picture. See figure 5 for an illustration of this sequence. After the event image had been presented for 5 seconds, a screen asking the subject to rate the agreement of the sentence and the picture appeared. The order of the sentence-picture pair was randomized by the E-prime software.

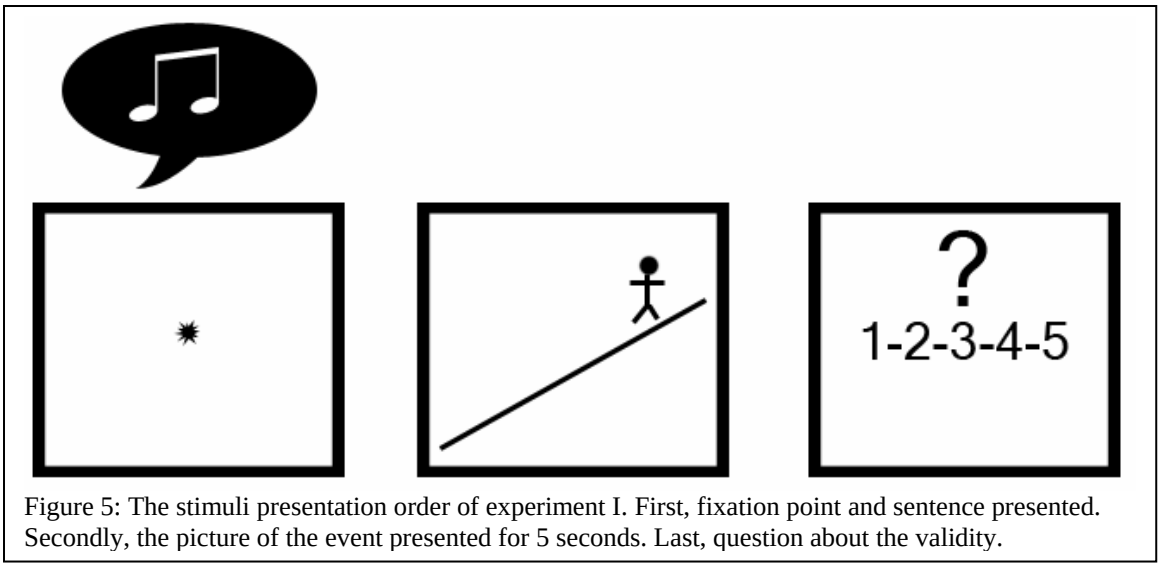

Figure 5: The stimuli presentation order of experiment I. First, fixation point and sentence presented. Secondly, the picture of the event presented for 5 seconds. Last, question about the validity.

7.4 Experiment II

7.4.1 Hypothesis

If the manner element changes the attention of the perceiver as described under experiment I, then the following is expected:

- The memory recall for details in the active zone should be significantly better for the group receiving sentences containing manner elements compared to the subjects receiving sentences without manner elements.

7.4.2 Procedure

After the eye-tracking experiment was finished, the subjects were asked if they could help out with a second, small experiment which was the memory recall test. If they agreed, which everyone did, they were led to another room where the computer running the recall test stood. The test consisted of pictures of the four motion event, but with some small alterations. The pictures of the events had been censored, namely the feet of the models. The subjects were presented with each picture and an accompanying statement, for example "Did the girl wear silver-colored shoes?". For two of the statements a 'yes' was the correct answer and for the

other two a ‘no’ was the correct answer. The subjects answered by clicking on an 11-step scale ranging from ‘yes’ to ‘uncertain’ to ‘no’, with intermediate steps to indicate the level of certainty/uncertainty.

7.5 Post-experiment procedure

After the experiment, general questions about the experiment were asked and the subjects got a chance to ask questions about the experiment. It was here the subjects were checked to see if they had second-guessed the aim of the experiment. The participants were also asked to read and, if they agreed, sign a paper of permission for the use of their data for academic purposes. The subjects received a paper giving a short explanation about the experiment (which, however, did not give away information on how to perform to confirm/disconfirm the hypothesis, should this information spread to other subjects). Finally, the subjects of the first test days received a bottle of wine as compensation for their time. The other subjects, who were tested in conjunction with another experiment, were not compensated. All in all, the experiment took about 20 minutes per subject to perform.

8. RESULTS

8.1 Experiment I

8.1.1 Manner versus non-manner: Event 1

In event 1, the picture showed a young woman walking/sliding down the hill. The manner sentences were “Tjejen glider ned för kullen” and “La fille descend la pente en glissant” (the girl slides down the hill”. The non-manner sentences were “Tjejen tar sig ned för kullen” and ”La fille descend la pente”. See figure 6.

Figure 6: Event 1. The active zone in green.

The average total fixation time of the target area for all subjects who received the non-manner version was 724 ms while the fixation time for the subjects who received the manner version was 1023 ms. These differences were not significant: $t(41) = 1.40$ ($p < .17$).

8.1.2 Manner versus non-manner: Event 2

In event 2, a young woman walking on a small road was depicted. The manner sentences were “Tjejen jäktar” and “La fille marche en haussant le pas” while the non-manner sentences were “Tjejen går” and “La fille marche”. See figure 7.

The average total fixation time for the target area for the non-manner subjects were 755 ms while the average total duration for the manner subjects were 990 ms. This was not a significant difference, $t(42) = 1.66$ ($p < .11$).

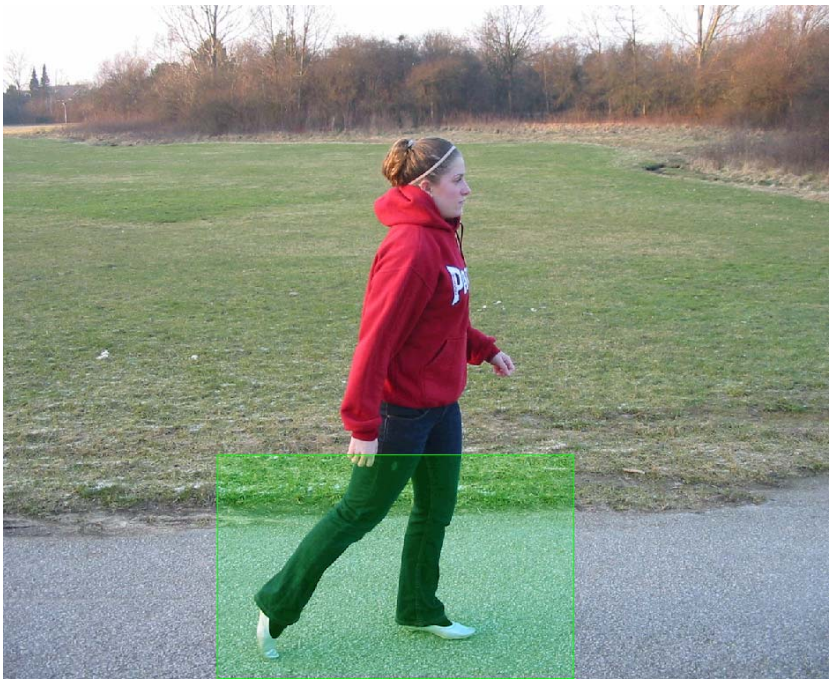

Figure 7: Event 2. The active zone in green.

8.1.3 Manner versus non-manner: Event 3

This picture showed a young woman running out of a house. The manner sentences were “Tjejen springer från huset” and ”La fille sort de l’immeuble en courant” whereas the non-manner sentences were ”Tjejen lämnar huset” and “La fille sort de l’immeuble”. See figure 8.

The fixation durations for the non-manner-tested subjects were on average 543 ms while they were 670 ms for the manner-tested subjects. The t-test shows that these differences are not significant: $t(42) = .84$ ($p < .41$).

Figure 8: Event 3. The active zone in green.

8.1.4 Manner versus non-manner: Event 4

The final event showed a young man sneaking into/entering a room. The manner sentences were “Den unge mannen smyger in i rummet” and ”Le jeune homme entre dans la chambre sur la point de pieds” whereas the non-manner sentences were “Den unge mannen tar sig in i rummet” and “Le jeune homme entre dans la chambre”. See figure 9.

In this last event, the average total fixation times for the non-manner subjects were 510 ms and 553 ms for the manner subjects. The differences were not significant at all: $t(42) = .37$ ($p < .72$).

Figure 9: Event 4. The active zone in green.

8.1.5 Manner versus non-manner: All events

No significant differences were found when looking at the events separately, but still, the manner group constantly had longer average total fixation durations in the target area for all the events. This calls for looking at all the events together. In order to compare scores between events, all scores were divided with the mean total fixation time for the non-manner group[3]. This means that the relative variation was preserved but the mean was moved in order to allow comparisons. The scores for the non-manner group were now positioned around 1 and the scores for the manner group were positioned around 1.27. These two new groups had their scores tested with the t-test[4], and the result was significant, $t(174) = 2.16$ ($p < .032$). Unfortunately, since all events needed to be merged to see an effect, it is not possible to conclude whether the semantic structure (as described in section 7.3.3) of the sentences help or hinder the effect.

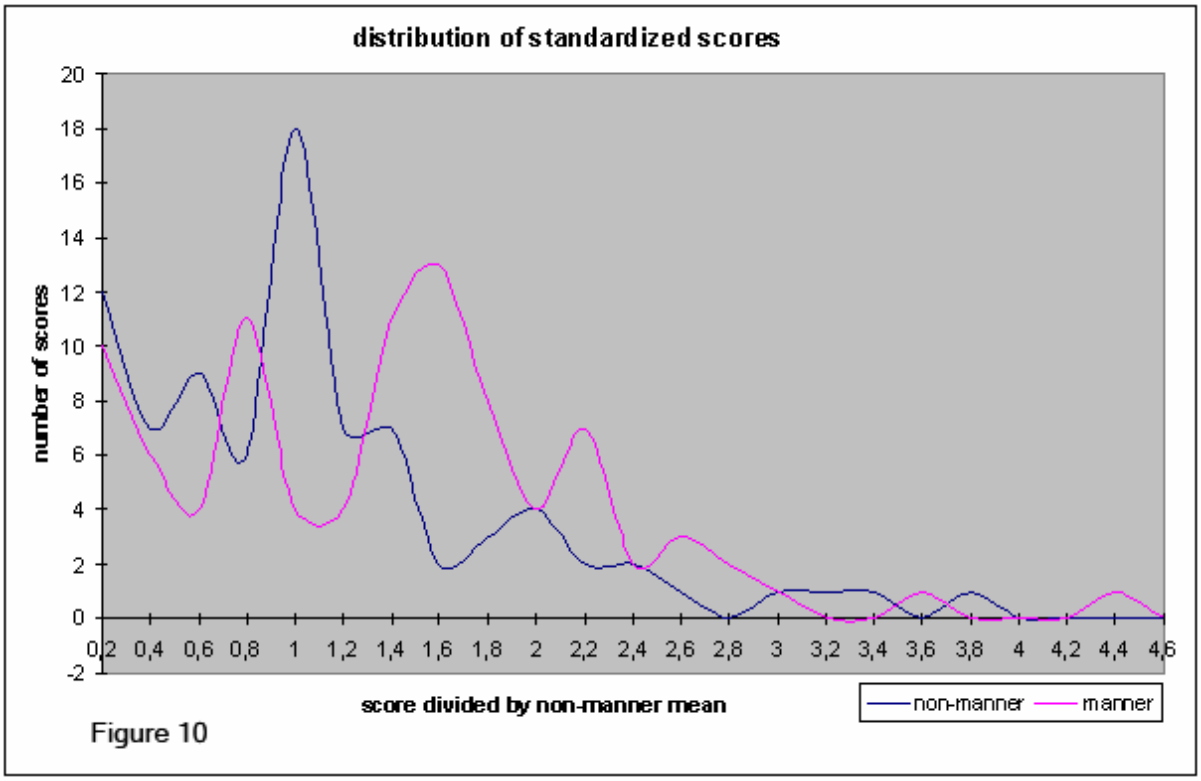

Figure 10

The complete distribution of the scores can be viewed in figure 10 above. We see that after the transformation of the non-manner scores they are positioned around 1 (the peak). The manner scores are then positioned in relation to the same mean used to transform the non-manner scores. The diagram shows that the manner scores are to a greater extent positioned to the right of the non-manner scores. Had the total fixation durations of the manners scores been generally lower than the non-manner scores, they would have been positioned to the left of the peak of the non-manner scores to a greater degree.

[3] This could also have been divided with the average for all subjects across groups but per event, or the average in the manner groups of each event. The important thing was that the scores were divided so their mean changed to allow for comparisons, but that their relation with regards to mean and variation to the other test group was preserved. To test that this transformation is valid I retested the t-tests for each event with their new numbers and the t-tests produced the exact same values. Furthermore, the statistical expert at SOL-centrum did not object to this transformation.

[4] A Kolmogorov-Smirnov test showed that the manner scores did not significantly deviate from normal distribution. The non-manner scores were not proven to be normal distributed. However, the sample size was large and the t-test is in this case robust for deviations from normal distribution.

8.1.6 French versus Swedish

Most of the test subjects are Swedish so it is very unlikely that we get significant results if we only look at the French subjects. However, since this thesis is about linguistic relativity, it is necessary to examine whether the French subjects behave as expected or deviate in an unpredictable manner. A significant difference between the French manner and non-manner group is unlikely since a merging of scores was necessary to get significant results for *all* subjects. Therefore, any test between two smaller groups are very unlikely to be significant. Simply a higher mean for the average total fixation duration is what we have to settle for in this comparison.

If we start by looking at event by event:

EVENT	FRENCH		SWEDISH		ALL	
	Non-manner	Manner	Non-manner	Manner	Non-manner	Manner
Event 1	1372 ms	1563 ms	521 ms	859 ms	724 ms	1043 ms
Event 2	796 ms	1053 ms	743 ms	967 ms	755 ms	990 ms
Event 3	556 ms	810 ms	539 ms	621 ms	543 ms	670 ms
Event 4	412 ms	560 ms	540 ms	551 ms	510 ms	553 ms
	(n = 5)	(n = 6)	(n = 16)	(n = 17)	(n = 21)	(n = 23)

Table 3: Manner and non-manner differences across languages

The data in table 3 reveal that French speakers behave much in the same way as Swedish speakers do, i.e. they look longer on the target area when presented with a manner sentence.

It also seems that in event 1 the French subjects looked considerable longer in the target area than the Swedish subjects. Indeed, a t-test reveals that both test variations, i.e. non-manner and manner, yields significant results between languages. That is, there are significant differences both between the French non-manner subjects versus Swedish non-manner subjects and the French manner subjects versus Swedish manner subjects. The t-tests showed $t(19) = 2.94$ ($p < .01$) for the non-manner comparison and $t(21) = 2.22$ ($p < .04$) for the manner comparison. It should be noted, though, that the French subject groups were not very large (n=5 and n=6), but these differences were not created by any statistical outliers. For French non-manner subjects, 4 out of 5 scores were over the Swedish mean and the French manner scores were all (6 out of 6) over the Swedish mean of the manner group. This result cannot be explained and is perhaps the effect of a variable outside the scope of this study

8.2 Experiment II

8.2.1 Manner versus non-manner: a memory recall test

Since results show that subjects receiving the manner version look longer at the target area they are expected to also remember parts of this area better. This memory test, however, reveals there are no significant differences in memory even with all manner and non-manner scores merged[5], $t(177) = 1.02$ ($p < .31$). The complete distribution can be seen in figure 11. This an unexpected result since previous research has shown a high correlation between eye fixations and memory. The complete distribution of these memory scores can be viewed in figure 11. In figure 11, the score 1 means the subject was completely certain that the incorrect word was correct. Score 6 means the subject was completely uncertain what was the correct answer. A score of 11 means that the subject was certain that the correct answer was in fact correct.

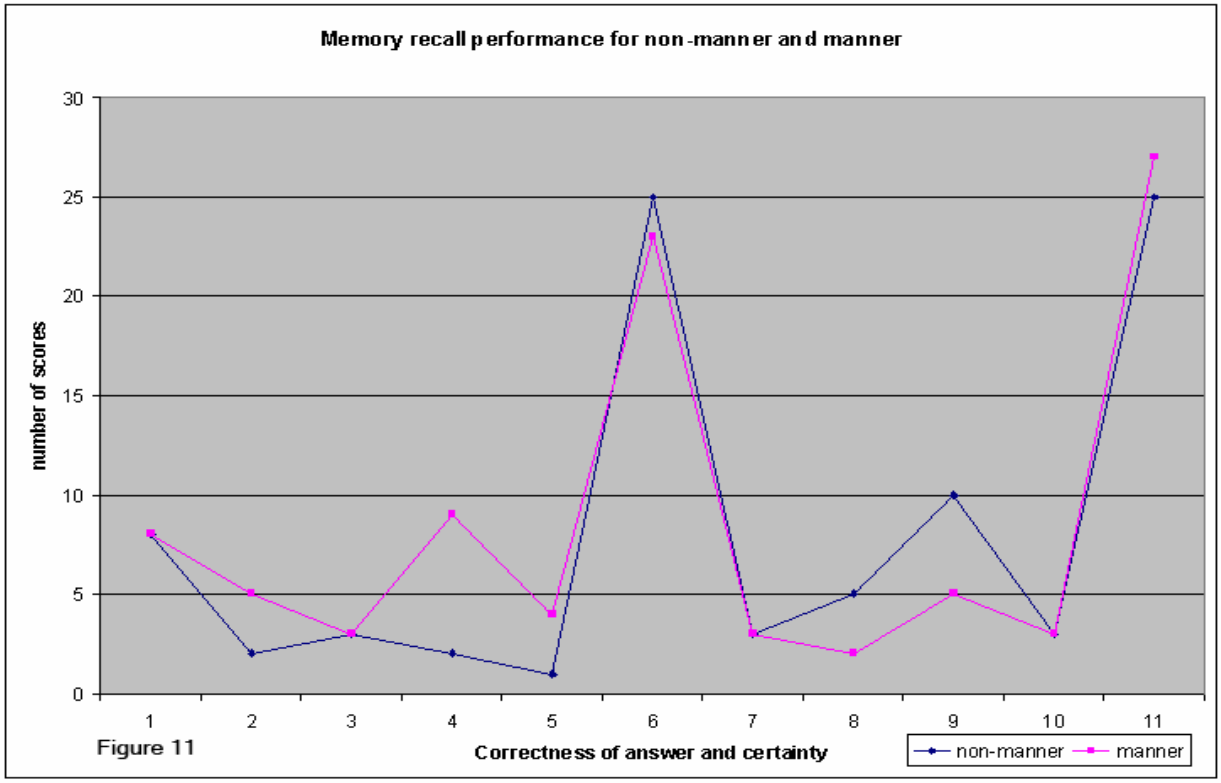

Figure 11

This diagram shows that the peaks of the two test versions are positioned around the same values. This clearly shows that there is no difference in memory between the two versions. The peak at 6 is the value for "uncertain/I don't know" and the peak at 11 is for a correct answer which the subjects were completely confident was the correct one.

[5] A Kolmogorov-Smirnov test showed that the two test groups were not normal distributed. However, it is believed the large size of the sample makes the t-test robust for this.

8.2.2 General high fixation total and memory recall

Because the manner and non-manner groups showed no differences in memory, it is called for to check whether the fixations in this test have an effect on memory at all. If fixation durations have an effect on memory, then the memory recall performance should be correlated with those subjects, from both the non-manner and manner groups, who have an above-average total fixation duration.

Four tests were made, varying events and whether the subject's fixation total was higher or lower than the mean (across both groups) in the target area of interest. Only event 4 was significant on itself, $t(42) = 3.58$ ($p < .001$), but if we look at the total merged scores, the results clearly show that fixations correlate strongly with memory recall[6]: $t(172) = 3.72$ ($p < .001$. The complete distribution can be seen in figure 12.

This figure shows that short fixation durations are mostly located in the middle of the scale (the "uncertain" area) whereas the long fixation durations are located in the far right, which is the area for a high level of correctness and certainty.

Even though the results both show that fixation durations correlate with memory and that the manner element correlates with longer fixation durations in the target area, no significant results linking the two are available. This will be discussed further in section 9.3.

The distribution of this recalculation of the recall test can be seen in figure 12.

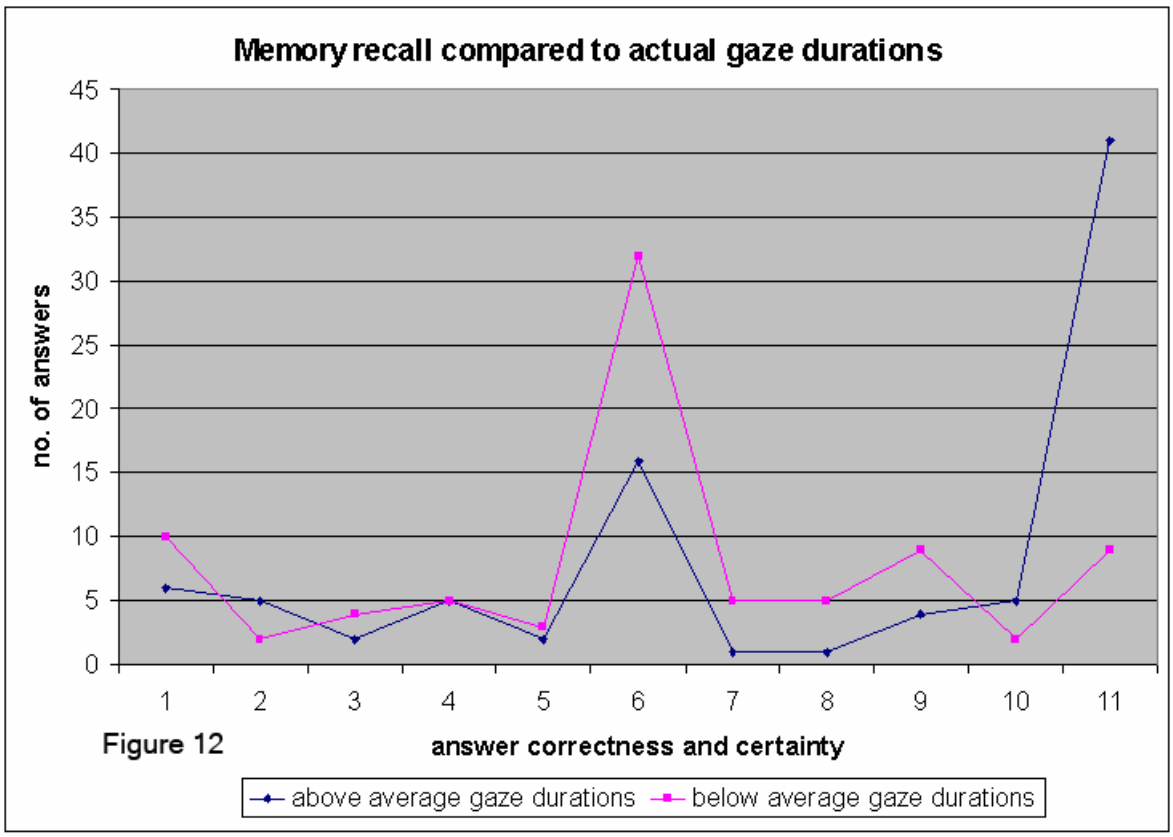

Figure 12

[6] Again, we rely on the large sample size and the robustness of the t-test for it to be an adequate test.

9. DISCUSSION

9.1 How the results relate to the hypotheses

The results show that the manner element failed to significantly direct the attention in the expected direction in any single event. However, if we standardize and merge the data across events, we get significant results. This indicates that the manner element affects attention in the previously mentioned way, but the effect is not strong. It also means that manner focuses attention to the active zones of the figure.

Furthermore, the fixations correlate well with memory recall, but the manner and non-manner groups do not. This will be discussed in detail in section 9.3 below. Several variables may be potential sources of errors or just generally need to be discussed before it is possible to conclude anything. After these discussions, we will return to the hypotheses.

9.2 The validity of the instructions

The subjects receiving the manner version of the test did look at the specified target area more. However, could this effect have been due to the instructions of the test?

The instructions stated that the subject would hear a spoken sentence, then look at a picture and finally determine how well the two fit together. Perhaps this yielded, under the five seconds each picture was presented, unnatural attentional behavior where the subjects would successively look at each identifiable semantic element presented in the spoken sentence to find a mismatch with the picture and that this is something they would not do in a naturalistic setting?

Indeed, this is something not controlled for in the experiment, but are these instructions unreasonable? I argue that we always operate under a certain set of "instructions" when we interact with other people. The spoken sentences are meant to simulate what another person is telling us and so the "instruction" we have is based on several contextual things. For example, if we think the person telling us this is boring then we do not really listen to what he/she has to say. The effect is that we pay little attention to the utterance and our viewing of the scene is carried out to a larger extent by bottom-up features such as movement, color, size, etc.

The context in which the instructions of this particular experiment would be the most ecologically valid, would probably be a situation where the listener is sceptical of the information conveyed by the speaker. This way, they would likely operate under an instruction which guided them to compare as much information as possible between the utterance and the scene to form a general opinion about the truth value of the utterance. However, the fact that subjects look at spoken counterparts (Tanenhaus *et al.* 2002) and semantic competitors with no overt spoken counterpart (Huettig & Altmann 2005), shows it is not unreasonable to assume the instructions in this experiment produced valid results.

If we are sceptical of the validity of the instructions, we can at least theoretically conclude that they are valid in some situations.

9.3 The validity of the memory recall test

Experiment II produced a surprising result, which did not fit with other results. First, experiment I proved that the manner element did attract attention to the active zone in the form of longer fixation durations. Second, fixation durations are correlated with memory recall. Thus, logically, the manner element would lead to better memory recall performance, but it does not. This is odd, but at least two explanations are possible. First, the manner element correlates with some other unknown variable which has a negative effect on recall performance. The second, more plausible explanation, is that it is due to statistical noise. The memory effect by the manner element is so small it did not become significant even though the recall scores from all events were merged together and tested as a whole. This is not unlikely given the fact that both the attentional effect of the manner element and the recall

effect of the fixations were significant only after the data of the events had been merged. This means that a combined effect would be even smaller and that the subjects are too few to show this. This could be illustrated as figure 13, where the initial effect gets smaller and smaller as it gets harder to distinguish from the noise for every new step it needs to move through. First, the subject has to pay attention to the manner element, which may not happen due to the subject having other thoughts. Secondly, the subject has to successfully encode the visual information located in the active zone into memory. Thirdly, the subject has to correctly recall the information. Every step introduces new noise which makes the effect harder to perceive.

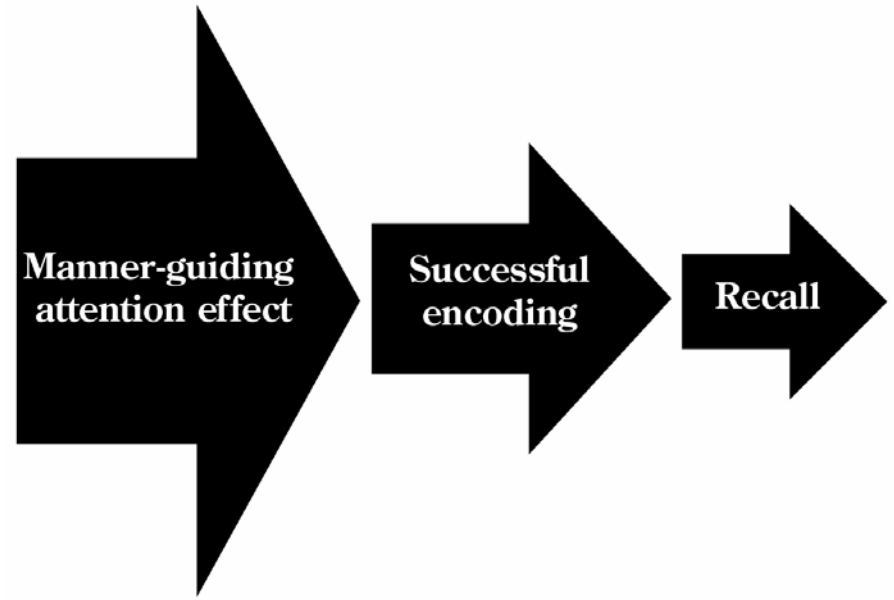

Figure 13: the decrease in signal strength of the examined effect.

9.4 The hypothesis revisited

Experiment I proves that manner elements guides the attention to the active zone, but that this effect is small. Experiment II, a memory recall experiment, shows that fixations are correlated with memory performance, but it failed to show a direct link between the manner element and the recall performance. However, it is very likely that this effect is there, but hidden in noise, since it should logically be present because manner correlate with fixations in the target area and those fixations correlate with memory performance.

10. CONCLUSION

The conclusion of this study is that the expected effect of linguistic relativity is there, but it is small. Depending on how sceptical one is of this study, the instructions, and of linguistic relativity in general, the perceived ecological validity will vary. At best, the effect studied here is valid at all times and produces an effect on attention of speakers every day which carries over to memory. At worst, the effect is limited to those situations where the listener is sceptical towards the speaker and the internal instructions resemble the sceptical ones used in the experiment and this limited effect is not carried over into memory. All in all, manner elements are likely to have an effect on cognition. Also, it has been shown that the attentional effect of manner elements guides the hearer to the manner's active zone of the figure

11. IMPLICATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Currently, research about speech-production and the link between prepositions and landmarks have begun in Humanistlaboratoriet. This study provides the means for conducting a similar production study investigating the link between manner and active zones. It would be interesting to examine this effect in a natural speech context such as a frog-story study. This study is also one of the first using eye-tracking to examine effects of relativity.

It would be interesting to repeat this study and use a larger selection of events which are systematically varied with regard to semantic structure. This way, it is possible to see when and how much the manner affects attention.

This empirical research concerning the attentional distribution of semantic elements is interesting because it gives an objective view on exactly what kind of informations is carried by the language and how much weight it is given. It also links abstract units such as semantic elements to concrete counterparts in the visual world.

REFERENCES

Alford, Danny, K. H. 1980. Part I: Demise of the Whorf Hypothesis. In *Phoenix: New Directions in the Study of Man.* Volume IV, Nos 1 and 2, 1980.

Allwood, J. 1983. "Kan man tänka oberoende av språk? in Teleman (Ed.) *Tal och Tanke*, Lund: Liber.

Bloom, P., Keil, F.C. 2001. “Thinking through language". Mind and Language, Vol 16-1, 351-367.

Brown, R. 1976. “Reference in memorial tribute to Eric Lenneberg”. In *Cognition*, vol 4, issue 2.

Cooper, R.M. 1974. “The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing”. *Cognitive Psychology, 6,* 84-107.

Corbetta, M. 1998. “A Common Network of Functional Areas for Attention and Eye Movements”. *Neuron*, Vol. 21, 761–773.

Feist, M. I. & Gentner, D. 2001. ”An influence of spatial language on recognition memory for spatial scenes”. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 279-284). Mahwah, NJ: Lawrence Erlbaum Associates.

Huettig, F. & Altmann, G. T. M. 2005. “Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm”. *Cognition*, vol 96, B23-B32.

Henderson, J. M. *et al*. 2003. “Eye movements and picture processing during recognition”. In *Perception & Psychophysics*, vol 65, 725-734.

Jacob, R.J.K., Karn, K.S. "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises”, (Section Commentary). In *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. Hyona, J., Radach, R., Deubel, H.(eds). 2003. Oxford: Elsevier Science.

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”. In *Journal of Memory and Language*, vol 48, 16-32.

Langacker, R. W. 2001. *Concept, Image and Symbol* – The cognitive basis of grammar. Berlin: Mouton de Gruyter.

Levinson, S. C. 2003. *Space in Language and Cognition* – Explorations in Cognitive Diversity. Cambridge: Cambridge UP.

Lucy, J. 1997. “Linguistic Relativity”. *Annu. Rev. Anthropol*. 1997-26: 291-312.

Lundqvist, D. and Holmqvist, K. (2004) “Bigger Is Better: How Size of Newspaper Advertisement and Reader Attitude Relate to Attention and Memory”, submitted.

Melcher, D., Kowler, E. 2001. “Visual scene memory and the guidance of saccadic eye movements”. In *Vision Research*, Vol 41, 3597-3611.

Nelson, W.W., Loftus G.R. 1978. “Cognition determinants of fixation location during picture viewing”. In *Journal of Experimental Psychology: Human Perception & Performance*, vol 4, 565-572.

Oh, K-J. 2003. Language, cognition, and development: Motion events in English and Korean. Unpublished doctoral dissertation. Department of Psychology, University of California, Berkeley.

Papafragou, A., Massey, C., Gleitman, Lila. 2002. “Shake, rattle ‘n’ roll: the representation of motion in language and cognition”. In *Cognition* Vol 84, 189-219.

Rayner, K. 1998. “Eye Movements in Reading and Information Processing:
20 Years of Research”. *Psychological Bulletin*, vol 124, No 3, 372-422.

Sigurd, B. & Zlatev, J. In press. *Festschrift for F.K.* Paper presented at the conference "Language, Culture and Cognition", July 16-18, Braga.

Slobin, D. 2003. “Language and Thought Online: cognitive consequences of linguistic relativity”. In Gentner, D. & Goldin-Meadow, S. (Eds.). 2003. *Language in mind: Advances in the study of language and thought* (pp. 157-192). Cambridge, MA: MIT Press.

Slobin, D. 2004. “The many ways to search for a frog: Linguistic typology and the expression of motion events”. In S. Strömqvist & L. Verhoeven (Eds.) *Relating events in narrative: Vol. 2: Typological and contextual perspectives* (pp. 219-257).

Slobin, D. In press. “Linguistic representations of motion events: What is signifier and what is signified?” In C. Maeder, O. Fischer, & W. Herlofsky (Eds.). 2005. *Iconicity Inside Out: Iconicity in Language and Literature*, Vol 4. Amsterdam/Philadelphia: John Benjamins.

Talmy. 2003. *Toward a Cognitive Semantics – Vol I*. Cambridge: MIT Press.

Talmy. 2003. *Toward a Cognitive Semantics – Vol II*. Cambridge: MIT Press.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., Chambers, C. 2000. “Eye Movements and Lexical Access in Spoken-Language Comprehension: Evaluating a Linking Hypothesis between Fixations and Linguistic Processing”. *Journal of Psycholinguistic Research*, Vol. 29, No 6, 557-580.

Whorf, B. 1956. *Language, Thought and Reality*. Cambridge: MIT Press.

Yarbus, A. L. 1967. *Eye Movements and Vision*. New York: Plenum Press.

Zlatev & David. 2004. Paper presented at the International Conference on Language, Culture and Mind, Portsmouth July, 17-20.

APPENDIX I
The event pictues and connected sentences in French and in Swedish:

Manner	**Non-manner**	**Manner**	**Non-manner**
La fille descend la pente en glissant	La fille descend la pente	La fille marche en haussant le pas	La fille marche
Tjejen glider ned för kullen	Tjejen tar sig ned för kullen	Tjejen jäktar	Tjejen går
			

Manner	**Non-manner**	**Manner**	**Non-manner**
La fille sort de l'immeuble en courant	La fille sort de l'immeuble	Le jeune homme entre dans la chambre sur la pointe des pieds	Le jeune homme entre dans la chambre
Tjejen springer från huset	Tjejen lämnar huset	Den unge mannen smyger in i rummet	Den unge mannen tar sig in i rummet
			

APPENDIX II

Swedish instructions for experiment I:

Du kommer nu att få höra en mening (på svenska) i högtalarna och därefter se en bild.
Din uppgift är att bedöma hur giltig meningen är i förhållande till bilden genom att ange ett värde mellan 1 till 5.
Ex:
1 = Ogiltig, hör ej samman med bilden.
3 = Vet ej/Osäker
5 = Fullständigt giltig, passar utmärkt väl ihop med bilden.

Detta kommer att upprepas 4 gånger.
Du svarar efter bilden visats färdigt.
Var vänlig och fokusera på stjärnan mellan varje bild (*).

(Tryck på MELLANSLAG för att börja testet)

French instructions for experiment I :

Vous allez à présent écouter une phrase et juste après une image s'affichera. Vous aurez alors à noter de 1 à 5 l'accord entre la phrase énoncée et l'image montrée.

Exemples :

1 = Désaccord, la phrase ne décrit pas du tout cette image.
3 = Incertain, difficile à juger.
5 = Parfaite adéquation, la phrase décrit exactement l'image.

Ce procédé sera répété quatre fois. Repondez après que l'image ait disparue. Veuillez concentrer votre regard sur l'asterisque entre chapue image.

(Appuyer sur la barre d'espace pour commencer le test)

English translation of the instructions:

You will now hear a sentence (in [language used]) from the speakers and then see a picture. Your task is to judge how valid this sentence is with regards to the picture by selecting a value between 1 and 5.
For example:
1 = Invalid, does not belong together with the picture.
3 = Don't know/uncertain
5= Perfectly valid, fits exceptionally well with the picture.

This will be repeated four times. You will answer after the picture has been shown. Please focus on the asterix (*) between pictures.

(Press SPACEBAR to begin the test)