

Non-referential Visual Contexts Effects on Incremental Semantic
Interpretation of Abstract Sentences: An Eye-Tracking Reading Study

by

Ernesto Guerra

A thesis submitted in partial fulfilment of the requirements for the degree of

Doktors der Philosophie

at the Faculty of Linguistics and Literature

Universität Bielefeld

April, 2014

Examination Committee: JProf. Dr. Pia Knoeferle at the Universität Bielefeld
PD. Dr. Falk Huettig at the Max Planck Institute for
Psycholinguistics
Prof. Dr. Stavros Skopeteas at the Universität Bielefeld
PD. Dr. Katharina Rohlfing at the Universität Bielefeld

ABSTRACT

This doctoral thesis focuses on two issues that, in the context of research into psycholinguistics, have been underexplored so far. First, we asked whether abstract language—which does not directly speak about objects and actions—could be affected rapidly and incrementally by visual information. Second, we ask whether visual contexts that are not a depiction of the object and events subsequently described in a sentence could affect processing time for that sentence. Additionally, this work aims to clarify the cognitive mechanisms that would allow rapid interaction between abstract language and what we denominated *non-referential* visual contexts.

A necessary literature review (Chapter 2) revealed a trend of moving from overt relationships between what is seen and what is described in a sentence, to subtler relationships, such as word associations. We identified however, that most studies examining real-time sentence processing concentrated on concrete language, and that results with respect to abstract language are mixed or limited to word recognition. Studies that were not particularly concerned with online sentence processing, on the other hand, have examined the relationship between visual information (e.g., spatial distance, containment and spatial locations) and abstract concepts. These studies covered a number of different conceptual tasks, such as similarity judgments, fast categorization, or ratings among others. They explored several abstract domains, such as the concepts of similarity, time, social relations and affective valence. None of those studies showed evidence of the involvement of spatial representations during online language comprehension.

We concentrated on two concepts, namely similarity and intimacy, for which we found compelling theoretical and empirical evidence suggesting a close link to spatial distance. We discussed existing behavioral evidence for this mapping and argue that it might have consequences for online abstract sentence processing. At the end of the chapter, we present

two existing models of situated sentence processing, discussing their central cognitive mechanisms and evaluating whether they could accommodate potential non-referential visual context effects in online abstract sentence processing. We argue that these accounts in their current versions are not well prepared to accommodate how visual scenes that do not depict the objects or events described in a sentence could potentially affect the processing of such sentence.

Using a novel eye-tracking experimental paradigm designed to investigate the four research questions outlined in Chapter 1, we conducted five experiments, in which we crossed two levels of spatial distance with two levels of abstract semantic content. Participants inspected a visual context with two playing cards that moved either close together or far apart. Immediately after, participants read a sentence that expressed semantic similarity between two abstract nouns (Chapter 3) or a social interaction between two characters (Chapter 4). Participants' eye-movements were recorded as they read critical sentences and subsequently reading measures were computed. The critical regions of the sentence were those in which similarity (either similar or dissimilar; Experiments 1, 2 and 3) or the type social relation (either friendly or unfriendly; Experiments 4 and 5) were implied or explicitly mentioned.

The results of the first three experiments revealed significant interaction between spatial distance and semantic similarity; overall, sentences that expressed similarity between two abstract nouns were read faster when playing cards in the visual context moved close together, compared to when they moved far apart. In contrast, sentences that expressed dissimilarity between two abstract nouns were read faster when playing cards in the visual context moved far apart, compared to close together. Analysis of the eye-tracking data in Experiment 4 further extended these findings showing that spatial distance can also modulate sentences concerning social relations. Interestingly, these results showed the opposite pattern to those experiments on similarity; sentences that expressed friendly social interactions between two characters were read faster when playing cards in the visual context moved far

apart, compared to when they moved close together, and vice versa for sentences expressing unfriendly social interactions. In Experiment 5, however, these effects were not replicated, suggesting that a fully non-referential visual context might modulate abstract language only in certain circumstances.

In the last chapter, we discussed the literature review chapter, summarizing key findings, open issues and specific research questions. We then discussed the results of our five experiments concentrating on the extent with to which they could answer our research questions. Subsequently, we contrasted our results with assumptions and predictions from existing models of situated language comprehension, and discussed some of the challenges that these findings presents for them. We argued that these accounts are not capable in their current state to accommodate our findings. However, we suggest that an extension of existing mechanisms such as co-indexing and spreading activation could account for our data. Consequently, we described how these updated mechanisms could explain the effects of a non-referential visual context on online abstract language comprehension. At the end of the chapter, we presented a summary of the key conclusions that could be drawn from this research project, and outlined a number of new directions in which our findings could potentially guide research concerning visually situated abstract language.

ACKNOWLEDGMENTS

I have been lucky enough to have the opportunity of following my dreams this far. I have been lucky enough to have met the people I met, to have been to the places I have been, to have heard what I heard and saw what I saw. This work has emerged from the inherent randomness of life. I have been an instrument. My ideas do not belong to me, for they are the result of an infinite ocean of coincidence in which I am just a drop of water.

Of the people I've met, I would specially like to thank my advisor, Pia Knoeferle for her guidance and support. I believe she found the perfect path between giving me enough freedom to explore my (sometimes seriously mad) ideas and keeping me focused, so I did not end up trying to track break-dancers' eyes as they spin and talk about philosophy. Much of this thesis is due to her trust and leadership. I also would like to thank the (former and current) members of the Language and Cognition Research Group at the Cognitive Interaction Technology Excellence Cluster, for providing a great working environment and vivid discussion during our colloquiums. Special thanks to Maria Nella Carminati, for always being there. I would also like to thank Marta Kutas for giving me the opportunity to be at her Lab and run my own (crazy) study, and for the intense discussion sessions we had together with Tom Urbach. Thanks for the challenge. Equally, thanks to the research-assistance team for proof-reading my German over and over. Special thanks to Jasmin Bernotat for helping me with data collection and many other (last minute) things and to Brigid Gallagher for her indispensable help in the last stage of completing my thesis. Without them, it would just not have been possible.

From those places I've been, I am extremely thankful to have had the chance to be part of the Graduate School of Cognitive Interaction Technology at the University of Bielefeld. It has been a very encouraging and inspiring environment. I am also grateful to have been part of the European Master in Clinical Linguistics where I had the luxury to study at the

University of Milano-Bicocca, University of Groningen and the University of Potsdam. I am convinced that this program led me to where I am now. Finally, to the institution in Chile to which I owe my most fundamental values, the Colegio Cahuala Insular my most heart-felt thanks for building an environment in which the truth and dialogue were more important than anything (anything) else.

Last but not least, I want to thank to my family, Claudia, Nelson, Francisca, Tatiana and Magdalena for their endless love, trust, encouragement and support. In the same way, this thesis would not have been possible without the companionship and love of my best friend and partner Johana. To many other unmentioned friends and colleagues this thesis is also dedicated.

TABLE OF CONTENTS

ABSTRACT.....	v
ACKNOWLEDGMENTS.....	viii
LIST OF TABLES.....	xiii
LIST OF FIGURES.....	xiv

Chapter 1: Introduction

1.1 A brief description of the state of the art.....	1
1.2 Research focus.....	3
1.3 Individual research objectives.....	5
1.4 Value of the research.....	5
1.5 Outline structure.....	6
1.5.1. Chapter 1.....	6
1.5.1. Chapter 2.....	7
1.5.1. Chapter 3.....	7
1.5.1. Chapter 4.....	8
1.5.1. Chapter 5.....	8

Chapter 2: From overt to subtle: a critical review of the interaction between visual information and sentence interpretation

2.1 Introduction.....	9
2.2. Language-modulated visual attention.....	10
2.2.1. Referential visual context studies.....	10
2.2.2. The role of working memory and different levels of representation.....	14
2.2.3. Summary and discussion.....	21
2.3. Language is grounded in direct experience.....	24
2.3.1. Perceptual representation and sentence processing.....	24
2.3.2. Perceptual representation and abstract concepts.....	27
2.3.3. SIMILARITY IS CLOSENESS.....	31
2.3.4. INTIMACY IS CLOSENESS.....	33
2.3.5. Summary and discussion.....	35

2.4.	Processing accounts of language in context.....	37
2.4.1.	The coordinated interplay account.....	39
2.4.2.	An affordance-based account.....	41
2.5.	Summary and conclusion.....	44

Chapter 3: Can spatial distance modulate incremental semantic interpretation?

3.1	Introduction.....	46
3.1.1	The present study.....	48
3.2.	Experiment 1.....	50
3.2.1.	Method.....	51
3.2.2.	Results.....	58
3.2.3.	Discussion.....	60
3.3.	Experiment 2.....	61
3.3.1.	Method.....	62
3.3.2.	Results.....	63
3.3.3.	Discussion.....	64
3.4.	Experiment 3.....	66
3.4.1.	Method.....	67
3.4.2.	Results.....	67
3.4.3.	Discussion.....	68
3.5.	Summary of results.....	70

Chapter 4: Can spatial distance modulate semantic interpretation of social relations incrementally?

4.1	Introduction.....	46
4.1.1	The present study.....	48
4.2.	Experiment 1.....	50
4.2.1.	Method.....	51
4.2.2.	Results.....	58
4.2.3.	Discussion.....	60
4.3.	Experiment 2.....	61
4.3.1.	Method.....	62
4.3.2.	Results.....	63
4.3.3.	Discussion.....	64

Chapter 5: Visually-situated abstract language processing

5.1	Introduction.....	89
5.2.	General discussion.....	91
5.2.1.	Literature review chapter.....	91
5.2.2.	Empirical chapter.....	94
5.2.3.	Interim summary.....	102
5.3.	Implication for models of situated language processing.....	103
5.3.1.	Visually situated abstract language comprehension.....	105
5.3.2.	A roadmap for processing accounts of visually situated (abstract) language.....	118
5.4.	Conclusions and future directions.....	122
5.4.1.	General conclusions.....	123
5.4.2.	Specific conclusions.....	124
5.4.3.	New directions.....	125
	German Summary.....	128
	APPENDIX	
	Critical sentences.....	134
	REFERENCES.....	146

LIST OF TABLES

3.1.	Summary of experimental conditions for Experiments 1, 2, and 3.....	53
3.2.	Mean reading times in milliseconds for Experiment 1.....	58
3.3.	Results from the linear mixed effects regression analysis for Experiment 1.....	59
3.4.	Mean reading times in milliseconds for Experiment 2.....	64
3.5.	Results from the linear mixed effects regression analysis for Experiment 2.....	65
3.6.	Mean reading times in milliseconds for Experiment 3.....	68
3.7.	Results from the linear mixed effects regression analysis for Experiment 3.....	69
4.1.	Summary experimental conditions for Experiments 4 and 5.....	76
4.2.	Mean reading times in milliseconds for Experiment 4.....	78
4.3.	Results from the linear mixed effects regression analysis for Experiment 4.....	79
4.4.	Mean reading times in milliseconds for Experiment 5.....	86
4.5.	Results from the linear mixed effects regression analysis for Experiment 5.....	86

LIST OF FIGURES

3.1. Levels of manipulation of spatial distance in the visual context.....	52
3.2. Schematic representation of an experimental trial.....	54
3.3. Log-transformed first-pass time for the adjective region in Experiment 1.....	60
4.1. Log-transformed first-pass time for the adverb region in Experiment 4.....	79

CHAPTER 1

Introduction

1.1. A brief description of the state of the art

In the context of psycholinguistics, the understanding of natural language processing is profoundly influenced by a robust empirical finding: visual referential context interacts rapidly with language processing (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Although previously studied by Roger Cooper (1974), it was not until Tanenhaus and colleagues (1995) showed that through a *referential link* non-linguistic context had rapid effects on syntactic structuring that this finding gained theoretical momentum. Tanenhaus et al.'s publication triggered a broad range of language processing studies using the *visual-world paradigm* (e.g., Allopenna, Magnuson & Tanenhaus, 1998; Altmann & Kamide, 1999; Chambers, Tanenhaus, Eberhard, Filip & Carlson, 2002; Dahan, Magnuson, Tanenhaus & Hogan, 2001; Kamide, Scheepers & Altmann, 2003; Knoeferle, Crocker, Scheepers & Pickering, 2005; Sedivy, Tanenhaus, Chambers & Carlson, 1999; Spivey, Tanenhaus, Eberhard & Sedivy, 2002; Tanenhaus et al 1995; Trueswell, Sekerina, Hill & Logrip, 1999).

Visual-world paradigms generally consist in the presentation of a visual context and linguistic inputs (e.g., a word, a sentence) that refer directly to those entities or events

depicted in the scene, while participants' eye movements are recorded. Since then, research has found evidence in support of an incremental and prediction-driven view of sentence interpretation, a process that can be rapidly modulated by non-linguistic visual information (see Altmann & Mirkovic, 2009).

Experimental evidence has shown that different aspects or features of visual context can interact with language interpretation, for instance the number of potential referents in the visual context (Tanenhaus et al., 1995; Spivey et al., 2002), visual referents' physical characteristics (e.g., size, shape) and action affordances (Chambers et al., 2002; Chambers, Tanenhaus & Magnuson, 2004; Dahan & Tanenhaus 2005; Sedivy et al., 1999), visually depicted actions (Knoeferle, Carminati, Abashidze & Essig, 2011; Knoeferle et al., 2005), speaker's gestures (Campana et al. 2005) and gaze (Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012; Staudte & Crocker, 2011). These studies have shown that non-linguistic and linguistic information is rapidly integrated when language refers to objects or actions in the visual context.

Some studies concentrated on how visual information affects early syntactic structuring and disambiguation (e.g., Knoeferle et al. 2005; Spivey et al. 2002), while others have focused on the relation between semantic clues gained in the visual context, and its effects on spoken-word recognition and prediction. Such studies have showed that the conceptual knowledge of an object can drive predictive eye movements (e.g., Altmann & Kamide, 1999; Kamide, Scheepers & Altmann, 2003) and that words can direct attention to unmentioned objects (e.g., Huettig & Altmann, 2005, 2011; Mirman & Magnuson, 2009). Furthermore, some studies have showed that language can direct visual attention to locations on a blank screen where visual referents were (Altmann 2004; Knoeferle & Crocker, 2007) or could be (Altmann & Kamide, 2009). Such results highlight the role of mental representation and working memory as mediators of the interaction between language and the visual world.

In sum, research into psycholinguistics has demonstrated that the interaction between language and visual information is (at least) two-folded; first, language interpretation can be informed by different aspects of visual contexts, and second, language can direct visual attention to objects in the visual context as they are (or are about to be) mentioned. Moreover, even when visual objects are not mentioned, they could attract preferential and predictive looks as a related word (e.g., Huettig & Altmann, 2005) or a linguistic context that enables prediction is processed (see Altmann & Kamide, 2007). Both effects seem to appear fairly rapidly and incrementally whenever there is a referential or a lexico-semantic link between what is seen and what is heard, and consequently in most cases when language is concrete (but see Duñabeitia et al., 2009). However, can more be said about the relation between visual perceptual processes and language processing? For instance, can non-linguistic visual information and language comprehension interact even without a referential or a lexical-semantic link? On the other hand, most concrete words have a clear referent in the real world, while abstract words do not unequivocally match a single picture. Thus, the other question that arises is whether abstract language comprehension interacts with visual information rapidly and incrementally.

1.2. Research focus

The present research project aims to address these two open questions. The first focus of this thesis is to investigate whether the interpretation of sentences formed of abstract nouns, adjectives and adverbs—which do not refer to particular objects in the visual context—could rapidly be modulated by visual information. The second focus of the thesis is to examine to what extent is a referential link or semantic association between objects and words is a condition for the interaction between language processing and the visual context. Existing studies in psycholinguistics have not yet addressed these two issues

comprehensively. Therefore, recent models of situated language comprehension have not explicitly considered how an apparently unrelated visual context could modulate abstract language comprehension. Consequently, the third focus of this thesis is to provide an explicit mechanism that could account for these potential effects.

A literature review informed our hypothesis with regards to relationships between abstract language and visual information. We found that on the one hand, psycholinguistic research had investigated a wide variety of world-language relationships, addressing an important number of relevant questions with regards to language processing and the interaction between linguistic and non-linguistic information. On the other hand, such studies have practically been limited to concrete nouns and verbs. Abstract nouns and adjectives have not, in this respect, been investigated in the context of online situated sentence processing. In contrast, an existing linguistic theory (conceptual metaphor theory, Lakoff & Johnson, 1980, 1999), presented non-experimental empirical evidence that suggests close connections between abstract concepts and perceptual representations. Additionally, our review of experimental approaches to conceptual metaphor revealed that the connections suggested by the theory also had implications for a number of offline and fast conceptual tasks, such as rating and similarity judgments. Yet, no evidence has shown that abstract language could be influenced by visual information during sentence processing.

To investigate these open issues we concentrated on two conceptual metaphors, namely SIMILARITY IS CLOSENESS and INTIMACY IS CLOSENESS (see Lakoff & Johnson, 1999). Existing experimental evidence showed how spatial distance and the abstract concepts of similarity and intimacy interact in various behavioural tasks. We discussed this literature in detail and argue that there is enough evidence suggesting that this kind of relationship—that has been under-explored from an online sentence processing approach—could be relevant during incremental language comprehension.

1.3. Individual research objectives

We conceptualized the identified gaps in the literature in two initial specific research questions (see below), which focused on the study of abstract language processing and visual contexts (I) and on the relation between visual context and language processing (II). Two additional specific research questions were addressed, aiming to understand the extent of the effects of the visual context and abstract language processing (III) and to specify potential mechanisms (IV) that seems to be relevant for the interaction between visual information and abstract sentence processing. A summary of these four research questions is presented here.

- I. Can abstract language interpretation interact with visual information in real-time, and if so, what is the time course of these effects?
- II. Can language-visual context interactions be observed in the absence of a referential or a lexical-semantic link?
- III. Does the effect of spatial distance on real-time semantic interpretation extend to different semantic domains, such as similarity and social relations?
- IV. Does co-indexing between abstract semantic interpretation and spatial distance depend on whether the linguistic element (i.e., the subject noun of the sentence) can be mapped in elements in the visual context?

1.4. Value of the research

The present research project addresses three important issues in the research of situated language comprehension. First, situated language processing research has been limited to the study of concrete nouns and verbs referring to depicted objects and actions, neglecting for the most part the study of abstract language comprehension¹. Thus, it is still

¹ It is worth noting that we acknowledge that psycholinguistics experiments have traditionally investigated a number of linguistic phenomena that can be considered to be quite abstract linguistic processes, for instance thematic-role assignment. Our focus is, however, different. We argue that the research on visually situated

unclear to what extent principles proposed with regards to the interaction between visual and linguistic processing can be generalized to situations in which language does not literally describe what is seen in the visual world. Secondly, studies that have investigated the relationship between abstract concepts and visual information have not yet addressed these questions from an incremental language approach. As a result, the time course of potential abstract language-visual context interactions remains unknown. In fact, it is not clear whether visual information would affect abstract language comprehension incrementally. Thirdly, we argue that existing accounts of situated language processing are not well equipped to accommodate potential findings that go beyond a referential or a lexical-semantic associative link. Accordingly, we suggest that it is necessary to revise those accounts, updating their mechanisms and some of their assumptions to accommodate visually situated abstract language processing. Addressing the three points described above is critical to understanding what kind of language-world relationships are relevant to language comprehension, what the time-course of these effects is, and what cognitive mechanisms make these effects possible.

1.5. Outline structure

1.5.1. Chapter 1: Introduction

The first chapter offers the reader a brief description of the state of the art of situated language comprehension research. It identifies a robust linking hypothesis in the study of language and visual context interactions, and the most commonly used methodology. Furthermore, it describes some of the questions that have been addressed, summarizing the most relevant findings, but also pointing to some open issues and questions that are still to be answered. The research focus is consequently outlined, and the specific research questions are

sentence processing has been limited to the study of visual contexts in which objects or actions are depicted, and to sentences that describe those objects and actions literally. In this sense, abstract language, such as abstract nouns (e.g., similarity, friendship) or abstract adjectives (e.g., similar, different, friendly, unfriendly), and their potential connections with visual information have not been assessed in the context of online sentence processing.

presented. Finally, the potential contributions of this research project are discussed, with a focus on the extension and update of current models of visually situated sentence processing.

1.5.2. Chapter 2: Literature review

Chapter 2 is divided into three sections. In the first section we review a large number of studies examining the interaction between visual information and language processing in real-time. We describe early and recent visual-world studies and the kind of questions that they have addressed moving from explicit referential links to subtler lexical associations relations between visual information and concrete language. The second section discusses literature from reaction-time studies investigating even subtler visual information-concrete language relationships. We then discuss a number of ratings and categorization studies that investigated the links between abstract concepts and visual information, although not during sentence processing. At the end of this section, we concentrate on examining evidence for two specific abstract concepts (i.e., similarity and intimacy) and argue that existing data suggests strong links between them and the spatial representation of distance, which could potentially be relevant for language comprehension. In the last section, we present the central mechanisms of two accounts of situated comprehension and discuss how they could be challenged by rapid and incremental effects of the visual context on abstract sentence processing.

1.5.3. Chapter 3: SIMILARITY IS CLOSENESS

In this chapter we present a novel eye-tracking paradigm, designed to examine potential effects of visual context that are not overtly related to subsequent target sentences. In three experiments, we examined the effects of spatial distance between objects in a visual context on reading times for sentences that expressed similarity between two abstract nouns. A detailed description of the experimental design, materials, procedure and data analyses is

provided. The results are presented in detail and briefly discussed. At the end of the chapter a summary describes the main results of the three experiments.

1.5.4. Chapter 4: INTIMACY IS CLOSENESS

Chapter 4 presents two further experiments intended to extend the results from the previous chapter. Using the same experimental paradigm, we examined whether the effects of spatial distance on semantic similarity interpretation observed in the first three experiments could extend to other abstract domains, namely social relations. As in the previous chapter, the experimental design, materials, procedure and data analyses, and the results are described in detail. A brief discussion section is presented for each experiment. At the end of the chapter there is a summary of the main findings from both experiments.

1.5.5. Chapter 5: Discussion

Chapter 5 is divided into three sections. First we present an in-depth discussion of our literature review chapter summarizing some of the key findings in the research of visually situated language. We focus this discussion on how we identified critical open issues and subsequently conceptualized them as research questions. Subsequently, we discussed the results from the five experiments presented in this thesis, concentrating on the extent to which they answer our research questions. In the second section, we contrasted our results with assumptions and predictions derived from existing models of situated language comprehension, and discussed some of the challenges that these findings present for them. We then discuss a number of cognitive mechanisms that could potentially account for our results. In the last section, we present a summary of the most relevant conclusions that could be extracted from this thesis and delineate some potential future directions of this research project.

CHAPTER 2

From overt to subtle: a critical review of the interaction between non-linguistic visual information and sentence interpretation

2. 1. Introduction

This chapter offers a review of the literature focused on the kind of visual context that informs language processing and the relationship between visual context and language that have been so far studied, mostly in the context of visually situated real-time language comprehension. We then review existing evidence on the effects of visual perceptual information (i.e., spatial information) on sentence processing and on abstract concepts. This review is not intended to be an extensive assessment of all available visual information-language studies rather it will focus on key articles that can exemplify a given phenomenon of interest. Its purpose, instead, is to summarize existing evidence about the interaction between language processing and visual contexts, identify potentially unexplored issues and generate a framework that can inform further research. Towards the end of the chapter, we will present two abstract concepts in which behavioral evidence suggests a clear connection to spatial information, however not yet in the context of incremental sentence interpretation.

Finally, we presented two existing accounts of situated sentence processing, namely the CIA (Coordinated Interplay Account, Knoeferle & Crocker, 2006, 2007) and an affordance-based account (Altmann & Kamide, 2007). These accounts offer a set of mechanisms that can account for a number of findings reported in the literature. We discussed those mechanisms and asked whether they would be sufficient to accommodate potential effects of visual context on abstract language processing and also whether they could accommodate about potential subtler relations between visual context and language processing. After the presentation of their main assumptions, we argued that those account are not well equipped to explain such (potential) effects.

2.2. Language-modulated visual attention

2.2.1. Referential visual context studies

After Tanenhaus et al. (1995) reported rapid visual context effects on syntactic structuring, research into psycholinguistics rapidly adopted the visual-world paradigm to investigate how language processing relates to visual perceptual processes (e.g., Allopenna et al., 1998; Altmann & Kamide, 1999; Knoeferle et al., 2005; Trueswell et al., 1999; for a review see Henderson & Ferreira, 2004). Generally, such visual-world paradigms examined how entities or visually depicted events inform real-time comprehension of spoken sentences that refer directly to those entities or events. Thus, in such experiments the visual referential context is intrinsically related to the linguistic information through a referential link.

Early studies focused on investigating different aspects of referentially-mediated visual context effects on syntactic processing. For instance, Spivey et al. (2002) examined the effects that the number of potential referents can have on processing locally ambiguous sentences. Syntactically ambiguous sentences present to the comprehender more than one

potential structural analysis. Thus, when participants heard an instruction such as “Put the apple on the towel in the box”, they could initially interpret the first prepositional phrase “on the towel” as the argument of the verb “put” rather than as a modifier of “the apple”. In their study, Spivey et al. argue that fixations to the incorrect goal-location (a towel) revealed that participants made such interpretation. Interestingly, this occurred when the visual context presented only one potential referent for the first noun phrase “the apple”. When the visual context presented two or more potential referents, the proportion of trials with a fixation to the incorrect goal-location decreased significantly. These results strongly suggest that the first prepositional phrase “on the towel” was interpreted as modifying the noun phrase “the apple” rather than the verb “put”. Thus, the number of referents in the visual context affected syntactic structuring, disambiguating the locally ambiguous sentence.

In another study, Chambers et al. (2004) asked whether knowledge about action affordances of objects (see Gibson, 1979) could also play a role in disambiguating locally ambiguous sentences such as “Pour the egg in the bowl over the flour”. Similarly, in Spivey et al.’s study, the first prepositional phrase of the sentence (“in the bowl”) could be interpreted as the argument of the verb “pour” rather than as modifying the noun phrase “the egg”. Participants were presented with a visual context with two potential referents for the noun “egg”, however, while in one condition both eggs were liquid, in the contrasting condition only one egg was. Thus, in one condition both potential referents could afford being poured, while in the other only one could be poured. When participants heard the locally ambiguous instruction while looking at the visual context with only one liquid egg (vs. two liquid eggs), a larger proportion of fixations to the incorrect goal-location (empty bowl) was observed. This effect was interpreted as reflecting a syntactic misinterpretation of the prepositional phrase “in the bowl” as the argument of the verb “pour”. The significant decrease in the proportion of fixations during the prepositional phrase when two visual

referents could afford being poured, suggested that knowledge about an object action affordance could modulate syntactic structuring incrementally.

World knowledge effects on the interpretation of unambiguous sentences have also been studied using the visual-world paradigm. In two experiments, Altmann and Kamide (1999) recorded participants' eye movements while these listened to sentences such as "The boy will eat/move the cake". The visual scene presented a boy and various objects including a cake, which was the only edible object. When participants heard "The boy will move the cake", the probability of fixations to the referent *cake* in the visual context increased after the post-verbal determiner onset and more so after the second noun phrase onset (particularly in Experiment 2). By contrast, when participants heard "The boy will eat the cake", the probability of fixations to the cake being the visual referent increased before determiner onset. Participants' gaze pattern could be interpreted as that they used the visually conveyed world knowledge (edibility of objects) to inform sentence interpretation in such a way that they could predict that the object that was about to be mentioned was the cake.

In a further study, Kamide et al. (2003) investigated the integration of syntactic information and world knowledge in predicting what is mentioned next. In Experiment 1, German native speakers listened to subject-verb-object (SVO) sentences such as "Der Hase frisst gleich den Kohl" (*The hare_{nom} eats shortly the cabbage_{acc}*) or object-verb-subject (OVS) sentences such as "Den Hasen frisst gleich der Fuchs" (*The hare_{acc} eats shortly the fox_{nom}*). In Experiment 2, native speakers of English listened to an English equivalent of the German SVO sentence ("The hare will eat the cabbage") or a passive version of the German OVS sentence ("The hare will be eaten by the fox"). While listening to such sentences, participants inspected a picture of a hare, a cabbage, a fox and a distractor object (a tree). The percentage of trials with fixations to the visual referent of the second noun phrase ("the fox") was calculated for regions prior to its acoustic onset. The results showed that the percentage

of trials with looks to the appropriate visual referent for the second noun phrase, was significantly higher right before “the fox” was enunciated (at the adverb “shortly” in the German experiment and the verb region in the English experiment). The authors concluded that during language comprehension both syntactic (case and voice) and semantic information (verb constraints), combined with real-world knowledge, can be used to make predictions about the most likely object to be subsequently mentioned.

In Altmann and Kamide and Kamide et al.’s studies, the visual context presented an arrangement of objects, but no actions. Furthermore, the linguistic input alone could have disambiguated thematic-role assignment before the determiner of the second noun phrase. To address these issues, Knoeferle and colleagues (2005) examined whether visually depicted actions alone could facilitate thematic role assignment incrementally. In their study (Experiments 1 and 2), participants inspected a visual scene with three characters, some of whom were depicted as performing an action upon another character. In critical trials, participants simultaneously listened to German SVO or OVS sentences (such as “Die Prinzessin wäscht offensichtlich den Pirat” ‘*The princess is apparently washing the pirate*’ or Die Prinzessin wäscht offensichtlich der Pirat ‘*The princess is apparently washed by the pirate*’, respectively). The locally ambiguous sentence was not linguistically disambiguated by the first noun “Die Prinzessin”, rather only after the the determiner of the second noun phrase (“den” vs. “der”). Thus, fixations to the appropriate object before the onset of the determiner would be strong evidence for visually driven syntactic disambiguation. Results from Experiments 1 and 2, support this hypothesis. The mean proportion of inspections to the appropriate role filler clearly increased after adverb (“offensichtlich”) onset; this was consolidated after the onset of the determiner of the second noun phrase. These results provided strong evidence that the action depicted in the visual context alone could facilitate thematic-role assignment.

The studies reviewed in section 1.1. are clear examples of how visual information can interact with language processing. Two insights can be clearly drawn from comprehenders' gaze patterns in these studies. First, they reveal that syntactic structuring can be updated moment-by-moment by extra-linguistic information, supporting an incremental view of sentence interpretation. Secondly, visual information (and semantic knowledge conveyed by it) constantly constrains and guides linguistic expectations, reflecting the prediction-driven nature of sentence interpretation. These effects can be observed whenever language is about the visual context, or in other words, when there is a clear referential link.

We now review a number of studies in which language directs listeners' visual attention to empty space or to objects that are not mentioned. This evidence will be discussed as regards the role which working memory and different levels of mental representations have on the interaction between language processing and the visual context. Further, we consider the implications of that evidence for the linking hypothesis between language and visual information.

2.2.2. The role of working memory and different levels of representation

As described above, Altmann and Kamide (1999) found that visual and linguistic semantic constraints allow listeners to look at objects in the visual context that were about to be mentioned. In a follow-up study, Altmann (2004) asked whether the co-presence of visual and linguistic information was necessary to observe such predictive saccades (see also Richardson & Spivey, 2000; Spivey & Geng, 2001). Similarly to Altmann and Kamide, participants inspected a visual context with two characters and several objects. However, in Altmann's study, the visual context was presented for five seconds and then replaced by a blank screen before the spoken sentence was presented. Eye-tracking data revealed that even

when the objects were no longer present, participants' looks to the appropriate location increased as they heard the verb of the sentence. These findings suggest that language-mediated visual attention is sensitive not only to co-presented visual information but also to the mental representation of such visual information, which proves to be true, within more or less the same time frame as when visual information is actually there.

In another study, Knoeferle and Crocker (2007) extended these findings to depicted events, and further refined the role of working memory in the interplay between language and visual scenes. In Experiment 1, participants inspected a scene (for seven seconds) with three characters and two depicted actions. The visual context was then removed before participants heard a related sentence. Crucially, the depicted actions were non-stereotypical for their agent (e.g., a *wizard* spying) while another character in the scene (e.g., a *spy*) was stereotypically related with such action. Target sentences could refer to one of the depicted actions (unique condition) or to a future action (ambiguous condition), which was semantically associated with one of the depicted characters (see also Knoeferle & Crocker, 2006). In the absence of the visual context, an increase in the proportion of inspections to the location where the correct agent was previously presented was observed even before it was enunciated. Results from Experiment 1 highlighted the priority of depicted actions (against world knowledge) for language comprehension, while replicating previous findings (Knoeferle & Crocker, 2006), even when the scene was removed before the spoken sentence began (cf. Altmann, 2004).

In Experiment 2, the characters remained on the screen and the depicted action was presented and then disappeared. The relative priority of the depicted action was observed for the unique condition, but no longer observed for the ambiguous condition: when the sentence was associated with both characters (stereotypically or by the depicted action), participants relied on their world knowledge (stereotypical vs. non-stereotypical characters) to establish reference before the character was actually mentioned. In Experiment 3, when both potential

referents were equally plausible (and thus world knowledge could not prompt participants to prefer either of them), the depicted action was again used to disambiguate reference. Based on these findings, Knoeferle and Crocker proposed that when visual information is not co-presented with language, its influence is subject to decay. In other words, whether a depicted event or other sources of information (e.g., world knowledge) are more relevant for language comprehension partially depends on how active such information is in working memory.

In another study, Altmann and Kamide (2009) investigated whether language-induced mental representations could guide visual attention to the potential location of an object in a visual scene. Participants were presented with a static scene co-presented with a spoken sentence (Experiment 1) or which disappeared before sentence presentation (Experiment 2). The scene presented a character and several objects. The critical manipulation was introduced through two sentential conditions: in one condition the sentence described the target object (the glass) as being moved from its location to a table-location (e.g., “The woman will put the glass on the table...”). In the other condition, the sentence described the glass-object as remaining at its location (e.g., “The woman was too lazy to put the glass on the table...”). Eye-movement patterns showed that as the sentence continued, the percentage of trials with fixations on the target-object (Experiment 1; or where it was, Experiment 2) decreased in the condition where language implied a change in location (vs. no change), while the opposite pattern was observed for the potential location (the table-object). These results suggest that the mental representation of the visual context can be modified by language, even when the unaltered visual context is co-presented with the linguistic input, but more so when the visual representation is only in memory.

Taken together, the results from the studies revised in this section show compelling evidence about one important aspect of the relation between visual referential contexts and language comprehension; rather than the simultaneous existence of a “visual world” as a

condition for the interaction between visual information and language interpretation, what seems to be crucial is the mental representation of such visual context (see Altmann, 2004). Moreover, the results show that such representations are subject to decay (Knoeferle & Crocker, 2007) and that they are constantly updated by linguistic information (Altmann & Kamide, 2009), among other sources.

A number of related studies provide further insights into how language interacts with different levels of mental representation. While previous studies showed that language can guide participants' looks to empty locations (where objects were or could be), we will now discuss how linguistic information prompts participants to look at objects that are not mentioned at all. In a visual-world study, Huettig and Altmann (2005) examined whether semantic features could drive listeners' visual attention to objects that, although not mentioned, share some of those semantic features with a target concept. Participants examined a visual context with four objects while a spoken sentence was presented. Three visual context conditions were compared, while the sentence remained the same. The visual context presented either the visual referent (e.g., a picture of a piano) of a target word and other unrelated distractors, or a semantically related picture (e.g., a picture of a trumpet) plus distractors, or both the referent and the related picture plus distractors. Eye-tracking data showed that the proportion of trials with fixations on the target object or on the semantically related object increased (compared to unrelated objects) upon hearing the target word, when the visual referent of the target, or the semantically related picture were presented. When both the referent and the related picture were presented simultaneously, a higher proportion of trials with fixations on the target object (vs. all other objects) were observed. In addition, the results showed a higher proportion of trials with fixations on the related object compared to the other two distractors. These results provide evidence that language can guide visual attention to objects that share semantic features, even when the objects are not mentioned.

In a related study Huettig & McQueen (2007) extended previous findings (Huettig & Altmann, 2005; see also Huettig & Altmann, 2007; Huettig, Quinlan, McDonald & Altmann, 2006) about the role of semantic features on language-mediated visual attention to other levels of representation. In Experiments 1 and 2, a visual context with four objects was presented to participants while they listened to a sentence. In relation to a target word, one picture was a phonological competitor (the name of the object shared the initial two or three phonemes with the target word; see also Allopenna et al., 1998; Dahan et al., 2001), one picture was semantically related to the target word (semantic competitor) and one picture had a similar shape to the visual referent of the target word (shape competitor). The fourth object was a distractor and did not hold any relation with the target word. Crucially, none of the spoken target words were visually depicted on critical trials. In Experiment 1, when participants could inspect the visual context for several seconds before the onset of the critical word, the results showed an early increase in the proportion of fixations to the phonological competitor, followed by an increase in the proportion of fixation to the semantic and shape competitors. This preference for phonological competitors rapidly decreased, while the greater proportion of fixations on the other two competitors was maintained. In Experiment 2, however, such early preference for phonological competitors was eliminated when participants were presented with the visual context only 200 ms before target onset. Instead, no differences were observed between the visual referent of the phonological competitor and the distractor picture. About 500 ms after word onset, a significantly higher proportion of fixations was observed (and maintained) for the semantic competitors relative to other objects. Finally, after 700 ms, shape competitors were preferred relative to phonological competitors and distractors.

These results showed that phonological representations have a relatively early advantage compared to semantic and visual representations, but this advantage depends on

whether the phonological representation has already been retrieved from the visual context upon hearing the target word (see Huettig & McQueen, 2007). When the visual context is presented only shortly before the target word (i.e., the names of pictures are not retrieved), this advantage disappears. By contrast, a short (Experiment 3) or long (Experiment 4) presentation of written words is sufficient to observe phonological advantage again. This suggests that while phonological information can be rapidly retrieved from written words, such information is less readily extracted from visual depictions. Moreover, these findings depart from a referential linking hypothesis since they showed that (word-associated) semantic knowledge can direct visual attention to visual objects even when no clear reference can be established. It is arguable that in such experimental context, preferred looks to unmentioned (but related) objects could be reflecting the attempt to find a “good-enough” object for reference establishment. Alternatively, these looks to unmentioned objects might reflect spreading activation of semantic and visual features from word recognition (see also Mirman & Magnuson, 2009).

Altmann & Kamide (1999) showed that based on verb-semantic and visual context constraints, comprehenders were able to predict the object most likely to be mentioned, as reflected by anticipatory eye-movements towards the appropriate visual referent. Huettig and colleagues (e.g., Huettig & Altmann, 2005; Huettig & McQueen, 2007; Huettig et al., 2006) showed that upon word recognition participants prefer to look at objects that share semantic features (e.g., semantic category, visual shape) with the referred expression. In a related study, Altmann & Kamide (2007) investigated whether they could observe anticipatory looks towards objects that were not mentioned, but that were semantically related to the referred expression. In two visual-world studies participants inspected a static scene with a character (e.g., a man), and several objects (e.g., a mug of beer, an empty wine glass and some cheese and crackers). The tense of the critical sentence was manipulated such that the participant

heard “the man will drink the beer” or the “the man has drunk the wine” (Experiment 1) or “the man will drink all of the beer” or the “the man has drunk all of the wine”.

The results from both experiments showed that at the onset of the verb complex (i.e., “will/has”) there was no bias towards any of the critical objects. At the onset of the verb, however, a larger proportion of trials showed fixations on the glass wine (in the past tense condition) and to the mug of beer (in the future tense condition). These results extend previous findings on anticipatory eye-movements (e.g., Altmann & Kamide, 1999) that reflected anticipation of a referring expression. Altmann & Kamide (2007) argue that the glass-object could not be expected to be mentioned since it would violate the verb constraints. Instead, comprehenders seem to be able to integrate world-knowledge, linguistic expressions and visual information rapidly such that they could predict that the sentence is about the empty glass of wine or the beer mug, before “beer” or “wine” are even mentioned.

Although all studies reviewed so far examined how visual attention is modulated by concrete sentences (concrete nouns and verbs), a more recent study investigated whether abstract language could also guide visual attention in predictable ways, and if so, what is the time course of these effects relative to concrete language. Duñabeitia et al. (2009) presented participants with four objects in a visual context and a spoken sentence simultaneously. The sentence-embedded target word was either the noun describing one of the pictures, an associated concept or an unrelated word. Critically, half of the associated and unrelated words were concrete and half were abstract. As in previous studies, Duñabeitia et al. found that upon hearing the target-associated word, the probability of fixations on the target object increased. This was true for both abstracts and concretes; however, participant’s eye gazes revealed that these effects occurred earlier for abstract compared to concrete target-associated words. These findings showed that both abstract and concrete words can guide visual attention to associated unmentioned objects, but with an advantage for abstract concepts.

2.2.3. Summary and discussion

Early visual-world studies exploited the evident referential link between a co-presented visual context and the linguistic input that describes it. In examining how language guided visual attention to empty space and unmentioned objects, the studies discussed in section 2.2.2. took a step away from the referential linking hypothesis (see Altmann & Kamide, 2007). Altmann (2004; see also Hoover & Richardson, 2008) demonstrated that the simultaneous presentation of visual information was not essential to the observation of rapid interaction between language and visual representations. Knoeferle and Crocker (2007) showed that previously presented visual information (and thus its mental representation) is subject to decay in working memory, such that participants can begin to rely on more salient sources of information instead (e.g., world-knowledge). Moreover, Altmann and Kamide (2009) showed that the mental representations initially built based on visual information can be constantly updated by language and world-knowledge. Other studies have showed that semantic (Huettig & Altmann, 2005; see also Mirman & Magnuson, 2009), visual (e.g., Dahan & Tanenhaus, 2005; Huettig & McQueen, 2007) and phonological (Dahan et al. 2001; Huettig & McQueen, 2007) similarities can guide looks to objects that are not mentioned, and that these sources of information have a predictable time course. Huettig and McQueen (2007) showed that during sentence-embedded word recognition, phonological information has a short-lived early advantage, while semantic and visual information require more time to begin to guide visual attention. Additionally, their results suggest that when the target-word is not visually depicted, pictures of objects of a similar shape gather more attention relative to semantically related objects, especially when presented shortly before the target-word. However, this effect disappears when provided with written words rather than pictures.

Together, this evidence highlights the importance of mental representation, in addition to visual information, as a mediator between language and visual attention. It also shows that

such representations are highly dynamic, since they can be quickly influenced by different sources of information, but they can also quickly decay and thus be less relevant for language interpretation under certain conditions. These results challenge a notion of language comprehension as a serial process (see e.g., Ferreira & Clifton, 1986; Friederici, 1995), and suggest instead that sentence interpretation is incremental and that sentence-embedded word recognition (or prediction) happens in cascade, where different types of representation are activated in parallel and mediate visual attention (see Huettig & McQueen, 2007). Indeed, existing evidence sheds light on the role that different levels of representation (i.e., phonological, semantic and visual) play on the interaction between concrete sentences and visual context. Similarly, a recent study (Duñabeitia et al. 2009) showed that abstract words can also guide visual attention in predictable ways through semantic association. However, virtually no studies have showed any interaction effects between visual representation and abstract language.

The revised literature shows evidently that concrete language (about objects and actions) has received substantial attention in the literature. By contrast, abstract language has been neglected by studies in psycholinguistics that use the visual-world paradigm. Indeed, abstract words can direct visual attention to objects that are associated with them (Duñabeitia et al., 2009). It is yet to be seen whether abstract language could rapidly interact with visual contexts beyond lexical-associations. Potentially abstract language could hold little or no relation to perceptual representations. In fact, Paivio (1986; see Crutch & Warrington, 2005 for empirical evidence) proposed a theory in which concrete and abstract concepts are represented in different ways; while abstract language is represented at the linguistic level and through associations with other words, concrete language is represented both at the linguistic and perceptual levels. Thus, it is possible that there are no relevant connections, beyond lexical-associations (Duñabeitia et al., 2009), between perceptual information and

abstract language. By contrast, concrete language could be rapidly related to visual perceptual information, since association between concrete nouns, for instance, and visual object recognition are vital during word learning (see Smith, 2009, in press).

Indeed there is mounting evidence for the interaction between concrete language and visual representations as discussed above. Meanwhile, such findings have been reported for abstract language only at the conceptual level (i.e., semantic associations). The lack of evidence, however, does not preclude the possibility that perceptual information can also play a role in the representation of abstract language. Alternatively, this gap in the literature regarding perceptual effects on abstract sentence interpretation could be related to a methodological bias towards concrete language in visual-world experiments. Initial studies that examined the influence of referential contexts on syntactic structuring (see Altmann & Steedman, 1988; Altmann, Garnham & Dennis, 1992; Ferreira & Clifton, 1986), did so using a written (referential) context, written target sentences and reading times as critical dependent measures. As noted by Spivey et al. (2002) such experiments have “used text not because the psycholinguistic community was primarily interested in reading *per se*, but rather because the theoretical questions required response measures that can provide fine-grained temporal information about ambiguity resolution” (p. 450). In this sense, the advent of the visual-world paradigm appeared to be an alternative to overcome some of the limitations intrinsic to reading times. Specifically, it elegantly allowed the study of the effects of non-linguistic referential contexts on language processing, while delivering a continuous measure that informed about visual attention moment-by-moment and time-locked to the incoming linguistic input.

Consequently, visual-world paradigms generally profited from the overt link between concrete words and their visual referents. But abstract language generally lacks a clear visual referent, indeed, can *freedom* be represented uniquely and undoubtedly by any image?

Whether this is the reason why psycholinguistics has not yet extensively studied abstract language or not, the fact remains that visual-world studies have not looked into the relationship between visual contexts and abstract language in detail. Paraphrasing Spivey and colleagues (2002: p. 450), visual-world paradigms have not only studied concrete language “because the psycholinguistic community was primarily interested in [concrete language] *per se*, but rather because [visual-world experiments] required [a referential link that could guide visual attention to objects when they are mentioned or are about to be mentioned]”.

As posited above, it is possible that abstract words are not represented at, nor associated with, the (visual) experience (see Paivio, 1986), and therefore, the investigation of abstract language and visual information is fruitless. However, there are several theories that propose that experience is relevant to all levels of abstraction in language, memory and thought (e.g., Barsalou, 1999, 2008; Glenberg, 2010; Lakoff & Johnson, 1999; Pecher & Zwaan, 2005; Zwaan, 2004). The next section revises some of the empirical evidence in support of such theoretical claims, and evaluates what can be extrapolated about incremental language comprehension from such evidence.

2.3. Language is grounded in direct experience

2.3.1. Perceptual representation and sentence processing

Theories of grounded cognition highlight the importance of experiential information for cognition, including language processing. They assume that mental representation largely rely on multimodal experiential traces (Barsalou, 1999, 2008). For example, the representation of an *apple* includes a stereotypical color, shape and weight, but also flavors, sounds and smells associated with the experience of interacting with an apple, in addition to the phonological and orthographic representation of the word “apple”. One framework

particularly concerned with the role of experience in language comprehension is the Situation Model (Zwaan & Radvansky, 1998; Zwaan, 2004). It proposes that during language comprehension experiential traces are activated through linguistic input, such that the comprehender constructs a mental model of the situation described. Consistently, Taylor and Zwaan (2008, see also Zwaan, Taylor & de Boer, 2010), proposed a linguistic focus hypothesis, according to which the activation of perceptual, motor and affective traces depends on how much the linguistic input focuses on those experiential traces. Although initially tested in the context of language-related motor resonance, namely the activation of motor cortical areas while comprehending language about actions (see Taylor & Zwaan, 2008; cf. Glenberg & Kaschak, 2002), a number of prior and later studies on the influence of visual information during sentence processing have shown to be consistent with this hypothesis.

For example, Zwaan and Taylor (2006, Experiment 5) asked participants to read sentences divided into segments of one to three words. Sentence segments were presented in the center of the screen, together with a visual stimulus consisting of 12 half ovals forming a circle around the location where words were presented. Such visual stimuli created an illusion of rotation. Critical sentences described a situation in which an act of rotation was performed (e.g., "... he | turned off | the | faucet"). Participants had to press the space bar to see each sentence segment. Analysis of the self-paced reading times revealed that a mismatch between the direction of the illusory rotation and the verbally implied rotation resulted in longer reading times. Crucially, this effect was significant at the verbal region ("turned off"), but not significant in previous or subsequent regions.

In a more recent study, Wassenburg and Zwaan (2010) examined whether the pre-presentation of pictures could modulate sentence reading times as a function of the match (vs. mismatch) between the orientation of the visual object, and the orientation of such object

implied in the sentence. Participants began by assessing 80 word-picture pairs in a verification task, out of which a number of 20 pictures were experimental items. In those pictures, an object was presented either with a vertical or a horizontal orientation. Finally and after performing a filler task (i.e., mental rotation task), participants were asked to complete a third task in which they read a total 40 sentences for comprehension. Each critical sentence named one of the objects previously presented in the experimental pictures of the first task, with a prepositional phrase implying that the object was either vertical or horizontal (e.g., "...the toothbrush in the sink/in the cup..."). In this way, critical sentences matched (or mismatched) with the pictures previously seen by the participants. During the reading task, participants' eye movements were recorded and two reading measures (first-pass and total reading times) were subsequently analyzed. Results showed that in the mismatch condition first-pass reading times were significantly longer at the prepositional phrase, compared to the match conditions. The same pattern was observed in total reading times at the prepositional phrase, however, only marginally significant.

The close time-lock of these effects supports the linguistic focus hypothesis and the importance of experiential traces in language comprehension. This agrees with evidence from visual-world experiments, which show that referential visual context effects arise when the relevant linguistic information directs visual attention to (or working memory representations of) visual features, which in turn inform language comprehension (see Knoeferle & Crocker, 2007). Moreover, paradigms such as those used by Zwaan and Taylor (2006) and Wassenburg and Zwaan (2010), showed that the influence of visual information on incremental sentence processing is observable in self-paced and eye-tracked reading times measures, and can provide not only information about difficulty of processing, but also important information about the timing with which visual context begins (and ceases) to be relevant (see Spivey et al. 2002 for a short discussion on the limitation of situated reading

studies). Finally, these studies showed that, just like language-mediated visual attention, sentence reading times can be modulated by visual information that is not directly referred to (cf. e.g., Huettig & Altmann, 2005; Huettig & McQueen, 2007).

Could these paradigms tell us anything about abstract language comprehension? Barsalou (2008, see also Zwaan & Madden, 2005) recognizes the challenge that abstract concepts represent for theories of grounded cognition, and makes an appeal to the role that perception of internal states might have for abstract conceptualization (see also Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). In this context, even from a grounded cognition perspective (which particularly concentrates on the role that experiential traces have on higher cognitive processing), it seems that visual representations have very little or nothing to do with abstract concepts, needless to say with incremental interpretations of abstract sentences. Some existing empirical evidence shows that comprehension of concrete and abstract verbs can modulate reaction times in visual detection tasks (Richardson, Spivey, Barsalou, & McRae, 2003), but elsewhere such results were replicated only for concrete verbs (Bergen, Lindsay, Matlock & Narayanan, 2007). To the best of our knowledge, there is no evidence showing that visual information can modulate incremental semantic interpretation of abstract sentences (see Pecher, Boot & van Dantzig, 2011 for a review).

2.3.2. Perceptual representation and abstract concepts

Linking visual perceptual processes with the interpretation of abstract concepts in language comprehension is more problematic compared to concrete concepts during language processing; abstract words and their relationships lack clear reference in the world. Nevertheless, Lakoff and Johnson (1980, 1999) developed a theory that directly related abstract language to experience. The authors observed that the link between abstract

conceptualization and experiential processing, although subtle in comparison to concrete concepts, could be seen even in everyday language. The conceptual metaphor theory (Lakoff & Johnson, 1999) is based on the analysis of recurrent linguistic expressions and according to it abstract concepts are metaphorically mapped in perceptual and motor representations. For instance, common linguistic expressions such as “Laura is my closest friend” and “She is such a warm person” are metaphorical in the sense that they are not intended to describe Laura literally. Instead, in saying “closest friend” and “warm person” what is actually meant is that Laura is “intimate” and an “affectionate person”. In accordance with theories of grounded cognition, conceptual metaphor theory suggests that this is not only a linguistic phenomenon but also a matter of cognitive conceptual structure (see also Gibbs & Colston, 1995 for discussion). Thus, highly abstract concepts such as intimacy and affection can be grounded, for instance, in spatial distance and temperature, respectively. Conceptual metaphor theory would describe these metaphors as INTIMACY IS CLOSENESS and AFFECTION IS WARMTH.

It would be inaccurate to say that there is no empirical evidence showing that (visual) perceptual processing can interact with abstract concepts. In fact, several behavioral studies have addressed the conceptual metaphor hypothesis, by exploring specific conceptual metaphors such as SIMILARITY IS CLOSENESS (Boot & Pecher, 2010; Casasanto, 2008) HAPPY IS UP (Meier & Robinson, 2004), GOODNESS IS BRIGHTNESS (Meier, Robinson & Clore 2004), CATEGORIES ARE CONTAINERS (Boot & Pecher, 2011) and TIME IS SPACE (Boroditsky, 2000; see Lakoff, Espenson & Schwartz, 1991 for a compilation of conceptual metaphors). However, none of these studies have yet looked into incremental language comprehension. We will move away from the language comprehension literature to discuss some examples of interaction between visual information and abstract concepts. For instance, a set of studies by Meier and colleagues (Meier & Robinson, 2004; Meier, Robinson & Clore

2004), explored the relation between visual information and emotional concepts.

In a first study, Meier et al, (2004) investigated the relationship between the brightness of a stimulus and its inferred emotional valence. In their experiments, positive or negative words were presented in black or white font against a gray background. The first three experiments differed only in the task emphasis (speed, accuracy, or both) and participants were instructed to categorize a given word in terms of its valence (positive vs. negative). The results showed longer reaction times (Experiment 2), lower accuracy rates (Experiment 3), or both (Experiment 1) for positive words in black (vs. white), and the opposite pattern for negative words. However, this pattern disappeared when instead of a categorization task participants were given a lexical decision task. In another study, Meier and Robinson (2004) investigated whether spatial information (e.g. word location in the vertical axis) could also have an effect on valence categorization of words. In experiment 1, participants had to decide if the word on the screen was negative or positive by pressing a button. A hundred words, half of them negative and half positive, were presented either at the top or the bottom of the screen. In addition, a spatial cue (+++) appeared before word presentation; first in the middle of the screen, then 1.5 inches from the center and finally, 3 inches from the center, marking the subsequent position of the word. The results showed faster response times for positive words when presented on the top (vs. the bottom) of the screen, and vice versa for negative words. In a subsequent experiment, the same set of words was presented in the middle of the screen for valence categorization, after which participants performed a visual-discrimination task on a cue presented either at the top or the bottom of the screen (cf. Bergen et al., 2007; Richardson et al. 2003). The results showed faster response times in the visual-discrimination tasks when word valence and cue position were congruent (i.e., positive-top, negative-bottom) compared to when valence and location were incongruent.

Other abstract domains such as time, categories, similarity and social relations have also received attention in the literature in experimental psychology (Boot & Pecher, 2010, 2011; Boroditsky, 2000; Casasanto & Boroditsky, 2008; Casasanto, 2008; Matthews & Matlock, 2011; McGlone & Harding, 1998; Núñez, Motz, & Teuscher, 2006; Williams & Bargh, 2008a). Evidence shows that visual spatial schemes, depicting ego- or object-movement can prompt participants to think about time in predictable ways, which then was reflected by their answer to a time relevant question (Boroditsky, 2000; see also Casasanto & Boroditsky, 2008). Another study, showed that when participants had to decide whether two visual objects belong to the same category, their response times and accuracy were modulated by the presentation of a frame that spatially contained (or separated) the two objects (Boot & Pecher, 2011).

We will now discuss two abstract concepts, namely *similarity* and *intimacy* for which non-experimental and experimental data converge in suggesting a strong link between them and the perceptual representation of spatial distance. Conceptual metaphor theory (Lakoff & Johnson, 1999) identified common linguistic expressions in which concrete language about space is used to express abstract concepts such as similarity and intimacy. According to Lakoff and Johnson, these expressions reflect the tight link that exists between abstract conceptualization and perceptual experience. More recently these connections have been examined from an experimental approach. Casasanto (2008) and Boot and Pecher (2010), investigated whether spatial distance could modulate similarity judgment ratings and response latencies. Other studies have asked whether spatial distance primes might affect people's judgment about their social relations (Williams & Bargh, 2008a), or whether language describing social relationships could prompt people to use space in predictable ways in a path-drawing task (Matthews & Matlock, 2011). A more detailed description of such studies and others is given in the next two sections.

2.3.3. SIMILARITY IS CLOSENESS

In everyday life, people usually refer to things that are similar as being close (e.g., “Your ideas are *close* to mine”), and describe things that are different as far apart (e.g., “My aims in this company are *far* from yours”). According to conceptual metaphors these expressions reflect a fundamental principle of the human conceptual system, namely that abstract concepts are grounded in perceptual and motor experience (e.g., Gallese & Lakoff, 2005). Initially, this claim was based on the observation of persistent use of spatial concepts when expressing abstract ideas (Lakoff & Johnson, 1980). More recently, experimental approaches have investigated whether those connections suggested by conceptual metaphors could also be reflected, in addition to everyday language, in behavioral measures, such as ratings and response times.

In a study (Casasanto, 2008), participants were asked to give similarity ratings (on a scale from 1 “not similar” to 9 “very similar”) for a pairs of stimuli, which were two abstract nouns (Experiment 1), two unknown faces (Experiment 2) or two line drawings (Experiment 3). In all these experiments each word/object were presented for two seconds, one after another and with a 500 ms blank screen in between. Participants then gave their rating with regards to how similar these two stimuli were. Critically, in the first two experiments the two stimuli were presented either as far apart, close to each other or at a middle distance on the computer screen. For Experiment 3, pictures were presented either close together or far apart. Results from experiment 1 (abstract word pairs), showed that distance between words affected participants’ similarity judgments; words shown close together were rated to be more similar than words shown further apart. In Experiment 2, the same paradigm was used but participants have to judge the similarity between two unknown faces. The results again showed a main effect of distance on similarity judgments. However, the effect observed was

the opposite to that observed in Experiment 1; faces presented close together were rated to be more different compared to words presented further apart.

In a follow-up experiment, Casasanto examined whether these conflicting patterns could be related to the type of judgment participants made. In Experiment 3, participants were presented with two line drawings, either far apart or close together for which participants had to perform a similarity judgment. Participants were randomly assigned to one of the two judgment conditions; half of them were instructed to perform a perceptual similarity judgment (“How similar in visual appearance?”), and the other half were instructed to perform a conceptual similarity judgment (“How similar in function or use?”) using the same scale from 1 to 9 as in Experiments 1 and 2. The results showed that distance affected similarity judgments distinctively depending on whether the judgment was perceptual or conceptual; when participants performed a conceptual judgment, objects presented close together received higher ratings compared to objects further apart, but when participants performed a perceptual judgment objects presented close together received lower ratings compared to objects further apart.

Another recent study also examined the conceptual metaphor theory prediction of the relationship between the abstract concept of similarity and spatial distance (Boot & Pecher, 2010). In a series of similarity-judgment reaction time experiments, participants were asked to identify, as quickly as possible, whether two squares were colored similarly or differently. Participants were presented with two colored squares on the computer screen. In Experiment 1, the stimuli were presented either close to each other, or far apart. In Experiment 2, a third middle-distance condition was included. Participants were instructed to press a yes-button when the squares were similarly colored and a no-button when they were different. The results showed that the spatial distance between the squares modulated response times for the similarity judgments; shorter response times were observed for squares with similar colors

when presented close together (as opposed to far apart), and shorter response times for squares of different colors when squares were far apart (as opposed to close to each other).

Findings as described by Casasanto (2008) and Boot and Pecher (2010; see also Breaux & Feist, 2008) imply that the relationship between spatial distance (perceptual) and similarity (abstract) can influence participants' decisions about how similar two stimuli are (given the graded rating scale), also how fast can participants identify whether two stimuli are similar or different (as reflected by response latencies). They show that the link identified by conceptual metaphor theory in common expressions in language (Lakoff & Johnson, 1999) can also be observed in conceptual tasks, such as judgments about similarity, for the SIMILARITY IS CLOSENESS metaphor.

2.3.4. INTIMACY IS CLOSENESS

As for the SIMILARITY IS CLOSENESS metaphor, recent behavioral studies have also investigated the relationship between the abstract notion of social bonds and perceived spatial distance. In a reaction time study, Bar-Anan, Liberman, Trope, and Algom (2007) examined whether the depth of a scene could interact with the meaning of words expressing social bonds (i.e., friend vs. enemy), in a distance-estimation task (Experiment 4) and a word classification task (Experiment 10). In their experiments, a set of pictures with a clear depth view (for example, scenery of alleys with trees) were presented together with a superimposed green arrow with a written word inside. The arrow was critically located in the picture in such a way that it seems to be near to (or far away from) the participant, due the depth perspective of the picture. Each trial presented either the word "friend" or the word "enemy". In Experiment 4, participants were asked to decide as fast as possible, whether the arrow was presented "near" or "far" from their perspective, and they were told that the written words

were irrelevant for the task. In Experiment 10, identical stimuli were presented; however, participants were instructed to classify the words by pressing a button whenever the word “friend” was presented and another button whenever “enemy” was the word presented. They were told that the depth of the pictures was irrelevant. Response times showed that in both tasks, perceived distance interacted with the content of the written words, congruence between visually perceived distance and social distance (namely, close-friend and far-enemy) produced shorter latencies, compared to the incongruent conditions (far-friend and close-enemy).

In another study, Williams and Bargh (2008a) asked whether visual spatial distance could affect participants’ judgments about the strength of the bond they felt for hometown and family. In a set of experiments, participants were instructed to mark off two coordinates on a Cartesian plane. Critically, participants were assigned to one of three coordinate distance conditions. The coordinates could have resulted in the two points being near to each other, in a middle distance, or far away from each other. Subsequently, In Experiment 4, they were given three questionnaires asking them to rate the strength of their bonds to their siblings, their parents, and their hometown, respectively. Participants rated their bonds on a scale from 1 to 7 (from “*not strong at all*” to “*extremely strong*”, respectively), after which the values from the three questions were averaged. The results showed that participants assigned to the far coordinates conditions, rated their bonds as significantly less strong compared to those participants assigned to the near coordinate conditions.

In a more recent study, Matthews and Matlock (2011) used a path-drawing task to examine the relationship between social bonds and spatial distance. Participants read a narrative describing a map scene with a clearly demarked path from a starting to a finishing point, crossed however, by barriers and figures symbolizing a character. The narrative talked about delivering a package by going through the park and passing different people. Critically,

the text ended by saying: “You know these people well. They are your friends.” (friend condition), or by saying “You do not know these people well. They are strangers.” (stranger condition), referring to the figures on the scene. Participants were instructed to draw the route between the starting and finish point in the park scene. Results from three experiments showed that the line drawn to mark the path through the park was significantly closer to the figures when described as friends, compared to when those figure were described as strangers.

Evidence from the above-mentioned behavioral studies reveal that the link between perceived spatial distance and social relations goes further than an everyday language phenomenon, and it can be observed experimentally in linguistic and non-linguistic tasks. Spatial distance can affect offline performance, such as rates about interpersonal relations and free-drawing of a path around “friends” (vs. strangers). Further, perceived distance in a scene can modulate reaction times in a word classification task, as a function of the social relationship given by context (i.e., “friend”, “enemy”), and similarly, word meaning can affect reaction times in a perceptual task (distance discrimination), depending on whether the word is presented visually nearer to or further from the participant’s perspective. These results are compatible with predictions that can be drawn from conceptual metaphor theory.

2.3.5. Summary and discussion

Evidence from studies such as those from Meier et al. (2004) and Meier and Robinson (2004) support the conceptual metaphor hypothesis about the relation between (visual) perceptual representations and abstract conceptualization of affective concepts. Interestingly, the visual feature critically manipulated (i.e., color, location) was irrelevant to the task (cf. Wassenburg & Zwaan, 2010; Zwaan & Taylor, 2006), and, moreover, in none of the studies the visual features were explicitly mentioned. In this sense, they suggest that subtler and more

implicit visual information (in contrast to hearing “apple” and seeing an apple) may well be relevant for conceptual processing, such as in a word-categorization task. Although, in principle this could suggest a certain level of automaticity, the effects of spatial location on reaction times disappear when participants performed a lexical decision task². Thus, while these results showed that low-level visual information (in this case spatial location, but also the color or shape of objects, or distance between them) can modulate reaction times in an abstract categorization task, it is unclear whether it is the result of strategic processing and whether they can be observed during language processing at all.

We review in more detail two abstract domains for which a number of behavioral studies have investigated their relationship to perceptual information. Evidence from a ratings study (Casasanto, 2008) and a reaction time study (Boot & Pecher, 2010) suggest that metaphorical mapping between similarity and spatial distance can be observed in conceptual tasks. Similarly, the results of from a reaction time study (Bar-Anan et al., 2007), a rating study (Williams & Bargh, 2008a) and a drawing study (Matthews & Matlock, 2011), showed that spatial distance can modulate the participant’s responses during word categorization, evaluation of the strength of social bonds, and the use of distance during path drawing, respectively. These results suggest that spatial distance is relevant to the conceptual representation of social relations and bonds. Online conceptual tasks (i.e., reaction times studies) and other offline tasks (i.e., ratings, drawing), however, do not provide evidence with regards to whether mapping abstract concepts and spatial distance could occur during real-time language comprehension, or about the time course of such potential effects. Addressing this question would require a method that provides a continuous measure which can monitor the potential effects of visual information on language interpretation moment-by-moment.

² In a lexical decision task, participants are asked to decide whether a string of letter is a word or not. Note that lexical decision tasks do not necessarily demand deep semantic processing (see James, 1975).

To summarize, while evidence coming from experimental approaches to conceptual metaphors (Boot & Pecher, 2010, 2011; Boroditsky, 2000; Casasanto & Boroditsky, 2008; Casasanto, 2008; Meier & Robinson, 2004; Meier et al., 2004) gives support to the metaphorical mapping hypothesis, none of the cited studies has looked at incremental language processing. Thus, it seems clear that visual information can modulate different abstract conceptual tasks when they involve offline decision making (Boroditsky, 2000; Casasanto, 2008) and also when they require rapid response (Boot & Pecher, 2010, 2011; Meier & Robinson, 2004; Meier et al., 2004). What remains unknown is whether the kind of link between abstract conceptualization and perceptual processing proposed by conceptual metaphor theory can be observed during online sentence processing.

2.4. Processing accounts of language in context

In the first section of this review we presented a number of visual-world studies in a descriptive manner, without much detail in regard to the mechanisms behind these effects. However, there are formal processing accounts, in the literature, that were designed to accommodate results such as those we discussed in the first sections. One of these accounts is the coordinated interplay account (CIA, Knoeferle & Crocker, 2006, 2007). In its current version, the CIA is able to explain how visual information, language and world knowledge interact rapidly, in contexts in which events are depicted and co-presented with spoken sentences, and also when events are presented not long before the spoken sentence. This account however, does not explicitly accommodate preferential or predictive looks towards unmentioned objects. In contrast another processing model, which we will refer to as an affordance-based account (Altmann and Kamide, 2007, 2009; Altmann & Mirkovic, 2009), does explicitly include, among their predictions, looks towards objects in the visual context that are not mentioned.

None of these accounts include specific predictions regarding abstract language, which is not surprising. The review above reveals that there are practically no studies investigating the relationship between visual context and abstract language. This means that up-to-date there are no findings to account for in this respect. Evidence revised in the second section however suggests that perceptual representations, in particular low-level visual information (e.g., location, distance, movement, etc.) can potentially be linked to abstract concepts as predicted by theories of grounding cognition (e.g., Barsalou, 1999; Lakoff & Johnson, 1999). It is yet to be seen whether those effects observed in off-line and fast responses to conceptual tasks (e.g., similarity judgments) can also be observed rapidly and incrementally during sentence comprehension. On the other hand, in those paradigms in which sentences guide participants' looks towards objects that are about to be mentioned or not mentioned at all, visual context still provides a set of objects that are or could be referred to in the sentence. In this sense, abstract language (since it does not refer unequivocally to physical objects in the real world) is a good case of study to investigate whether a fully non-referential visual context—that is, a visual context containing objects that are neither referred to nor lexically-associated with the content of the target sentence (for instance, two playing cards)—could in fact modulate incremental language processing.

In summary, there is no evidence that visual information could interact with abstract language processing incrementally, and additionally, it is unclear whether a non-referential visual context could affect language comprehension. Perhaps, abstract language does not relate to visual scenes after all and there is nothing more to be said about the relationship between visual information and language processing. Likewise, it is possible that although it has been shown that people prefer to look at objects related to linguistic inputs, even if unmentioned, a fully non-referential visual context might lack the linking hypothesis necessary to connect linguistic and non-linguistic information. However, if these two

hypotheses were correct, processing accounts of language in context will need to integrate such findings into their models. But could they already? In the following section we will describe the basic mechanisms and structures of the above-mentioned accounts focusing on whether, and if so how, they could accommodate effects of a non-referential visual context on abstract language comprehension. Most of the findings, for which these models account, were described in the first section of this Chapter. Thus, we will mainly concentrate on the mechanisms that explain those findings, referring back to specific results when necessary.

2.4.1. The coordinated interplay account

The CIA (Knoeferle & Crocker, 2006, 2007) is a formalized processing model which is aimed at accommodating the interaction between visually depicted events, stored knowledge, visual and internal attention and language comprehension. Broadly speaking, it assumes that this interplay occurs in three stages, however the authors explicitly embrace the possibility that these processes could partially overlap and happen in parallel, depending on which source of information is more accessible at a given point in time. In the first stage, comprehenders interpret linguistic input incrementally including expectations associated with it. In the second stage, partial sentence interpretation directs (visual) attention to relevant aspects of the visual scene (or its mental representation in working memory) through a referential mechanism and/or through expectations derived from the linguistic input. The model assumes that the working memory buffer is subject to decay, which means that the mental representation of a previously inspected scene might decay and thus become accessible to language comprehension as time goes by. In the final stage, linguistic interpretation and the active non-linguistic representation resulting from the visual context are co-indexed and the interpretation is revised if necessary, or in other words is grounded in the visual scene. This tightly time-locked coordinated interplay (the above-mentioned three stages) begins

again when further linguistic information becomes available.

In its current version, the CIA can accommodate results such as the relative priority of depicted events for language comprehension (Knoeferle & Crocker, 2007). As mentioned above, depicted events have priority only if they are still available for language comprehension (depicted or in working memory), and as long as other sources of information, such as world-knowledge, do not compete for attention in the system. For this, the role of working memory appears to be critical in the account (for a description of the results see pp. 15-16). Another aspect of this account that appears to be essential is the referential link between visual context and language. The following quote expresses the essential role that this relationship plays on the model;

“The CIA is an account of situated utterance comprehension. In situations where the utterance does not directly relate to the immediate visual environment, immediately depicted events will almost certainly not have the importance that is suggested by findings from Experiments 1 and 2 in this article because the scene is irrelevant. It is the immediate presence and relevance of utterance, linguistic/world knowledge, and depicted events for comprehension that enables the rapid interplay between these informational sources. We do expect, however, that in situations where the utterance is about the immediate environment, our findings of the rapid, verb-mediated influence of depicted events on structural disambiguation, and of the priority of depicted events over verb-based thematic role knowledge in thematic interpretation will apply.” (Knoeferle & Crocker, 2006: p. 526).

and in the revised version of the account, the authors further wrote;

“Naturally, however, not every object mentioned in the utterance will be in the scene, and often utterances will not—or will only tangentially—relate to the immediate scene/environment. The Coordinated Interplay Account assumes that the relative importance of scene information is a function of the success of referential processing throughout the utterance.” (Knoeferle & Crocker, 2007: p 520).

In these two paragraphs it is made clear that there must be an explicit relationship between the content of an utterance and the visual environment. In fact, the second stage proposed in the CIA relies on a referential or anticipatory search mechanism, towards an

object expected to be referred to. Thus, attention is directly mediated by the utterance in so far as that utterance is about the visual context. If effects of non-referential visual context are observed rapidly and incrementally during sentence processing, the CIA in its current version will not be sufficient to accommodate them. Without an attentional mechanism that could guide (visual or internal) attention towards relevant aspects of the visual context, co-indexing between different sources of information cannot occur, and therefore non-linguistic information could not be made available for language comprehension.

Is it possible that co-indexing occurred before establishing a reference and independently of it? The original and the revised versions of the CIA emphasize that the three steps that constitute the key processes are presented in quasi-linear order for the sake of simplicity, and in fact they can partially overlap and happen in parallel. This means that potentially, co-indexing could work (to a certain extent) independently of attentional mechanisms. The co-indexing mechanism, however, is specified to map a limited number of representations that are similar in language and the real world, namely nouns with objects and verbs with action. It is therefore a challenge for the CIA to accommodate the potential effects of a visual scene on abstract nouns, such as “war” or “peace”, since they do not refer to a specific object, or abstract adjectives, such as “similar” or “friendly”, since they refer to a relational concept or an abstract property of an entity, respectively.

2.4.2. An affordance-based account

An affordance-based account has been proposed by Altmann and Kamide (2007), which explicitly aims to accommodate preferential and anticipatory looks to objects in the visual context that are not mentioned at all. A central idea of this account is that conceptual-feature overlap underlies drift of attention towards objects that are not mentioned. The idea of

conceptual overlap is based on a notion that has been influential for language processing models, which is that similar or related representations are activated in parallel (e.g., McClelland & Rogers, 2003; McClelland & Rumelhart, 1981, Rogers & McClelland, 2008). Such connectionist approaches to cognition assumed that semantic representations (or concepts) are graded structures composed by parts or features, which in turn are shared with other concepts (see also Cree & McRae, 2003; McRae, Cree, Seidenberg & McNorgan, 2005). In this account, if certain features are activated by a (visual or linguistic) stimulus—say, is-musical-instrument for the word “piano” as in Huettig and Altmann (2005), concepts that share this feature (e.g., “trumpet”) would be more readily activated. In this sense, conceptual overlap results in a boost to the activation of related representations. The authors further argue that this mechanism is not theoretically different from priming effects explained in terms of spreading activation (see Collins & Loftus, 1975).

Conceptual overlap thus explains why comprehenders are more likely to look at the visual depiction of a trumpet when hearing the word “piano” (Huettig & Altmann, 2005, see also Dahan & Tanenhaus, 2005, and Huettig & McQueen, 2007, for visual features overlap; Myung, Blumstein, & Sedivy, 2006 for action-affordance overlap). But how does this account explain anticipatory looks to objects that are not going to be mentioned? Affordances at this point become critical to the account. As in the above-mentioned example, the word “piano” causes an activation boost of other concepts such as trumpet, because they share a conceptual feature—that is, both the piano and the trumpet afford to be played. In the case of predictive eye-movements towards objects that are about to be mentioned, as in Altmann and Kamide (1999) “the boy will eat... (the cake)” experiment, comprehenders anticipate that the object that affords to be eaten (the cake) is going to be mentioned (see also Kamide et al. 2003). Finally, in the case of anticipatory fixations on objects that will not be mentioned, as in Altmann and Kamide (2007) “the man will drink/has drunk... (the beer/the wine)”

experiment, an empty glass is the object that is more likely to afford the containment of wine that in turn has already been drunk (in the case of past tense), while the mug full of beer is the object in the visual context that is more likely to afford the containment of beer that can be drunk. To this end, Altmann and Kamide (2007) suggest that the conceptual fit of mapping between language and the visual context happens at the level of affordances.

In contrast to the CIA, an affordance-based account is rather underspecified with regards to the links between the visual context and language that are necessary for the rapid interaction of language processing and visual information. Rather, Altmann & Kamide's account has a general mechanism, namely the conceptual overlap of affordances, which guides all drifts of attention in the cognitive system through spreading activation of shared features. Indeed, this account offers an explanation beyond a referential linking hypothesis. But can this account accommodate potential effects of fully non-referential visual context? The following paragraph helps to clarify this question;

“Sentences such as *‘the boy will eat the cake’* are dynamically unfolding representations of events. The scenes that have typically been used in our corresponding visual world studies have simply been static representations of the objects that can take part in these events— they depicted *states*. [...] when mapping sentences onto static scenes, the system must determine which part of the denoted *event structure* the scene depicts; it must determine whether the scene depicts the start state, the end state, or an intermediary state. Thus, the process of interpreting a sentence, in the context of mapping that sentence onto a scene, requires the interpretation of that scene with respect to the event structures (and their temporal properties) entailed by the sentence. [...] studies in which participants are expected to map the event representation associated with an unfolding sentence onto that associated with a (semi-realistic) visual scene and the events that it affords do require this aspect of temporal interpretation (see also Knoeferle & Crocker, 2006). It is this aspect of the process that will allow us to explore the representational content that drives language-mediated anticipatory eye movements.” (p. 504, emphasis from the original).

As can be interpreted from the above paragraph, an affordance-based account requires that comprehenders map the interpretation of the scene and the interpretation of the sentence.

Although less explicitly than in the CIA, this means that the visual context must depict at least a potential state (if not the full event structure, see Knoeferle & Crocker, 2006), which in turn allows comprehenders to link the derived visual representations to the interpretation of the unfolding sentence. Thus, it is not clear how a fully non-referential visual context could be mapped to the event described in the sentence, in order to make predictions with respect to forthcoming linguistic information, when the visual context does not depict the events described by sentence in the first place. As for the CIA, it seems that the affordance-based account requires further specification in order to explain potential effects of non-referential visual contexts.

Finally, the affordance-based account seems to be able to accommodate recent results on the interaction of abstract language and visual information. Findings such as those from Duñabeitia et al. (2009), can be explained in terms of spreading activation from abstract words such as “smell” towards associated concrete concepts (e.g., nose), resulting in preferential looks towards the visual depictions of a related words. The affordance-based account is greatly underspecified, which in principle enables it to explain this kind of abstract language effect so far, at least, at the lexical level. This effect could be characterized as semantic priming and depend specifically on the activation of lexical associations, and do not require compositional processing as the anticipatory effects described. More importantly, it is not sensitive to contextual appropriateness (Huettig & Altmann, 2007).

2.5. Summary and conclusion

In summary, this review presented a discussion of a number of phenomena studied in the context of situated language comprehension. It concentrates on the type of visual world-language relationships that allow the interaction between linguistic and non-linguistic

processing. Specifically, the review was intended to show how studies moved from very explicit relationships to increasingly subtler ones. To that extend, we asked whether abstract language could also be visually situated. The review revealed the lack of evidence that language is sensitive to visual scenes beyond a referential link or lexical-semantic associations. Yet, results from studies less concerned with the time course with which visual information can modulate abstract language have shown evidence that suggest potentially interesting subtler world-language relationships.

We shall argue that there is enough theoretical and empirical evidence suggesting that visual information can interact with language comprehension, beyond a referential link or lexical-semantic associations. However, psycholinguistic research has poorly examined this possibility, and therefore, little is known about the time course of these potential effects. A discussion on existing processing accounts of situated language comprehension revealed that they are not well prepared, in their current versions, to accommodate potential effects of fully non-referential visual contexts. Visual information could affect early abstract language processing (e.g., 250 ms after word onset or after first fixation), or alternatively at later stages (e.g., end-of-sentence response times, total reading times) as the critical words are integrated into the linguistic context of the sentence. By using continuous measures (eye-tracking reading measures), it is possible to explore this question.

CHAPTER 3

Can spatial distance modulate incremental semantic interpretation?³

3.1. Introduction

As discussed in Chapter 2 evidence from research into situated language processing has demonstrated that non-linguistic contextual information can be used during sentence comprehension when language is about the objects in that visual context, a claim largely supported by empirical evidence (e.g., Altmann & Kamide, 1999, 2009; Chambers, Tanenhaus, & Magnuson, 2004; Knoeferle & Crocker, 2006; Tanenhaus, Magnuson, Dahan & Chambers, 2000; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Spivey, Tanenhaus, Eberhard & Sedivy, 2002). Visual-world studies have found robust evidence as to how the interaction between linguistic and non-linguistic information drives people's visual attention in predictable ways. Specifically, the fact that people look spontaneously at objects in the visual context as they are mentioned in a sentence has been interpreted as reflecting processes of establishing a reference.

But language can direct people's attention to unmentioned objects in a visual context, as well (e.g., Huettig & Altmann, 2005, 2011; Huettig & McQueen, 2007; Huettig & Hartsuiker,

³ The data presented in this Chapter have been also submitted for publication in the form of a journal article.

2008). A number of visual-world studies have more recently examined how objects in a visual context, semantically related to a (target) word, attract more spontaneous gazes toward them compared to objects that are not semantically related. Systematic looks to objects that are not directly referred to, have been interpreted as reflecting co-activation of phonological, semantic or perceptual information during word- and picture-recognition (see Huettig and McQueen, 2007). On the other hand, these spontaneous looks to unmentioned objects could also reflect attempts to establish reference and thus ground language to the available visual information (see Altmann & Kamide, 2009 for a similar notion).

Previous visual world experiments showed that phonological competitors (words that begin with the same phonemes as the target) strongly compete for attention especially during early stages of establishing reference (see Dahan, Magnuson, Tanenhaus & Hogan, 2001; Huettig & McQueen, 2007). Objects that are semantically related to target words are also stronger competitors compared to non-related competitors, and receive a higher percentage of looks upon hearing the target word (e.g., Huettig, Quinlan, McDonald & Altmann, 2006; Mirman & Magnuson, 2009). Finally, objects that share visual characteristics such as color or shape with a target referent have also been found to attract more visual attention than objects that do not share these visual characteristics (see Huettig & Altmann, 2007, 2011). These results suggest that a referential link between language and visual context is not a condition for interaction between non-linguistic and linguistic information.

Existing visual-world paradigms, therefore, directly exploit a referential or a lexical-semantic associative link between linguistic and non-linguistic information. In the first case, there is (more or less) a one-to-one relation between what is seen and what is heard. Thus, participants are effortlessly able to map linguistic information onto visual information (e.g.,

Kamide et al., 2003; Knoeferle et al., 2005), perhaps as a way to ground (or in other words directly connect) language with experience. In contrast, when participants are presented with unmentioned objects (e.g., Huettig et al., 2006), they systematically prefer to look at objects related to the linguistic target input. What is still an open question is whether non-linguistic information that is not referred to, and that is not associated through lexical-semantic links, can rapidly modulate incremental language interpretation. In a set of visually situated reading experiments we addressed this issue.

3.1.1 The present study

A visual world paradigm with spoken sentences would not adequately inform our question, since we are looking to a subtle (non-referential) relationship between language and visual information. However, recent studies have shown that visual context can influence reading predictably, when language refers to the visual context (Knoeferle, Urbach & Kutas, 2011), and also when critical visual information is incidental (Wassenburg & Zwaan, 2010). Thus we implemented three eye-tracking experiments, where we presented a visual context followed by a written sentence. The visual context presented two objects (i.e., two playing cards). Knowing that visual context can rapidly interact with language comprehension when they hold a direct relationship (for instance, a referential relationship: seeing a picture of an apple and subsequently reading a sentence using the word “apple”), as the experiments proceeded we gradually moved away from this relationship until the visual context was not overtly related to the following sentence. In Experiment 1, visual context was related to the sentence through written words on the cards for critical items. In Experiment 2, as in Experiment 1, half of the cards in the visual

context presented words on them. This time however, this was true only for fillers and never for critical trials. In Experiment 3, all cards in the visual context were empty and did not advance any words before the sentence. Thus, there was no overt relationship between the visual context and the subsequent sentence. It is worth noticing, however, that the critical manipulation in all the experiments was not based on whether the visual context was related or not (via words on cards) with the following sentences, but rather on the spatial information conveyed by the distance between cards in the visual context.

Based on a metaphorical mapping mechanism (Lakoff & Johnson, 1999), clear predictions can be made with respect to the connections between perceptual experience (e.g., observable distance between objects) and abstract language (e.g., semantic similarity). In contrast, a metaphorical mapping mechanism does not make any predictions about the time course of potential effects during real-time sentence comprehension. Recent work however, suggests that information from different modalities is immediately and incrementally activated during reading (e.g., Hagoort & van Berkum, 2007; Knoeferle, Urbach & Kutas, 2011; Wassenburg & Zwaan 2010). Thus, if perceptual information such as distance between objects is part and parcel of the semantic interpretation of similarity, we should see immediate effects of distance in a sentence reading task (e.g., first-pass reading times). Current situated language processing accounts (e.g., Altmann & Kamide, 2007; Knoeferle & Crocker, 2006, 2007) don't explicitly include mappings of the kind we have been discussing. They recognize the role of non-linguistic contextual information for language interpretation, relying however, mostly on a referential link between visual and linguistic information. To the extent that we find (potentially rapid) effects of spatial distance on semantic similarity interpretation during comprehension, situated language processing accounts will want to accommodate them more explicitly.

3.2. Experiment 1

Current research in psycholinguistics has not yet examined the extent to which visually derived perceptual representations, namely spatial distance between objects, could modulate semantic interpretation of abstract sentences. It is also unclear the extent to which visual context can influence incremental language comprehension in the absence of a direct referential or lexical-semantic associative link. We conceptualized these two open issues as two research questions: (a) can the semantic interpretation of abstract sentences be rapidly and incrementally modulated by non-linguistic information? and (b) can non-linguistic visual information modulate language comprehension even in the absence of referential or lexical-semantic links? This study intends to address these two open issues by examining the effects of spatial information from the visual context in the comprehension of abstract sentences.

A recent similarity-judgment study reported that when written abstract words were close to one another, they were judged as more semantically similar, compared to when the same words were presented far apart (Casasanto, 2008, Experiment 1). In our first study we examined whether this spatial distance effect extends to incremental comprehension during sentence reading. Participants saw two cards depicted as moving either far apart or close together, and which presented either semantically similar or dissimilar abstract nouns. They then read a sentence that contained these two nouns and judged sentence veracity based on their world knowledge. If spatial distance effects generalize from similarity judgments to sentence reading, we should find shorter reading times for sentences conveying similarity and dissimilarity between the two abstract nouns when the preceding words-on-cards were close to each other and far apart, respectively. Moreover, if spatial distance affects semantic interpretation incrementally, these effects should emerge at the adjective region of the sentences, since it is at this point in the

sentence that the relationship between the two words is made explicit (either similar or dissimilar). Note that, in principle, spatial distance effects could emerge even earlier, at the second noun phrase, since the *and*-coordination of the first two sentential nouns already implied either semantic similarity or dissimilarity. Observing these effects in early reading time measures (first-pass time, regression path duration), would suggest they are rapid; alternatively, spatial distance effects could appear in later measures (e.g., total times) or later (post-adjective) in the sentence.

3.2.1. Method

3.2.1.1. Participants

Thirty-two native German speakers (mean age: 23,6; range of age: 19 to 33) with normal or corrected-to-normal vision from the Bielefeld University community received €6 each for participating. None of them had been exposed to a second language before age 6. All participants gave informed consent.

3.2.1.2. Materials and design

The visual context for experimental trials showed the images of two playing cards, each sized 155 x 265 pixels. They had an identical blue playing card design on one side, and a solid grey color on the other side. Each of them had an abstract noun (e.g., *peace*, *war*) written in black 12-pts. font on its front side (see Figure 2.1.), both of which appeared in the subsequent sentence. For experimental trials, we manipulated the distance between these cards (and thus between the words on them), so that experimental visual context had two levels; in the *CLOSE* level, the two cards moved from their initial position closer to each other. In the *FAR* level, the

two cards moved from their initial position further from each other. Before constructing the sentences for the eye-tracking experiments, we conducted a similarity judgment pre-test. Participants ($n=16$), rated a large set of preselected pairs of abstract nouns on a scale from 1 to 7. This list included 60 pairs of synonyms and 120 pairs of antonyms, plus 30 related and 30 unrelated pairs of words as fillers. Participant from the rating study did not take part in any of the eye-tracking experiments. From these we selected the 48 noun pairs with the highest inter-rater consensus⁴ and constructed coordinated noun phrase sentences that either expressed similarity (e.g., *'Battle and war are surely similar, suggested the anthropologist'*) or dissimilarity (e.g., *'Peace and war are certainly different, suggested the anthropologist'*). Words that differed between the sentences within an item (the first noun phrase, the adverb and the adjective) were matched for number of characters and frequency⁵.

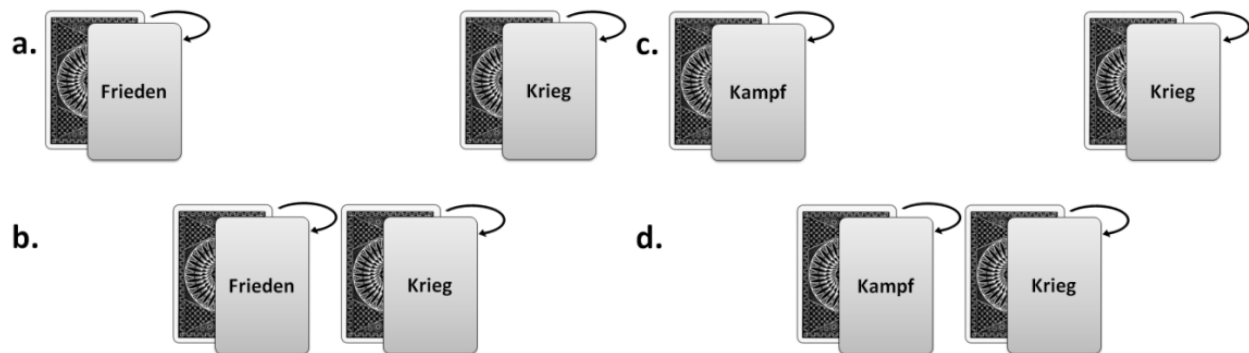


Figure 3.1. Example of visual contexts for the sentences in Table 1. Cards appeared on the screen showing their back and then turned, as indicated by the arrow, to present the two first sentential nouns in Experiment 1.

⁴ Item 48 in all three experiments and item 47 in Experiments 1 and 3, were removed from the analysis due to an error in the order of presentation of the words (see Appendix A).

⁵ Two paired sample t -test revealed no significant differences between the frequencies or the length of the two groups of words (both t -values (48) < 2, p -values > .05).

A 2x2 within-subjects within-items experimental design was used. The combination of the two experimental visual context levels (*CLOSE* or *FAR*) and the two experimental sentence levels (*SIMILAR* or *DISSIMILAR*) resulted in four experimental conditions, which were presented within participants. Thus, participants saw cards close to each other followed by a sentence that expressed similarity (Close-Similar); cards close to each other followed by a sentence that expressed difference (Close-Dissimilar); cards far apart followed by a sentence that expressed similarity (Far-Similar); and cards far apart followed by a sentence that expressed difference (Far-Dissimilar) as describe in Table 3.1.

Table 3.1: Example item sentences and the four experimental conditions

Image	Sentence	Condition
Fig.1a	Frieden _{NP1} und _{coord.} Krieg _{NP2} sind _{VP1} bestimmt _{ADV} verschieden _{ADJ} , das verriet _{VP2} der Anthropologe _{NP3} .	Far-Dissimilar
Fig.1b	Frieden _{NP1} und _{coord.} Krieg _{NP2} sind _{VP1} bestimmt _{ADV} verschieden _{ADJ} , das verriet _{VP2} der Anthropologe _{NP3} .	Close-Dissimilar
Fig.1c	Kampf _{NP1} und _{coord.} Krieg _{NP2} sind _{VP1} freilich _{ADV} entsprechend _{ADJ} , das verriet _{VP2} der Anthropologe _{NP3} .	Far-Similar
Fig.1d	Kampf _{NP1} und _{coord.} Krieg _{NP2} sind _{VP1} freilich _{ADV} entsprechend _{ADJ} , das verriet _{VP2} der Anthropologe _{NP3} .	Close-Similar

Translation. Dissimilar sentence: ‘Peace and war are certainly different, suggested the anthropologist’;
Similar sentence: ‘Battle and war are surely similar, suggested the anthropologist’.

To assign items and conditions to participants, we used a Latin square design: experimental items were divided into four lists, each of which contained the same number of trials per condition and every item in only one condition. In addition to the 48 items each list also contained 96 filler trials (144 trials in total); 24 of the fillers displayed nouns on the front side of

the cards in the visual context, while 72 had blank fronts. Card presentation before the sentence was similar to the experimental trials varying either in their final position (e.g., top corners vs. bottom corners) or their color (e.g., red cards). Trial presentation was pseudo-randomized such that all experimental sentences were either preceded by one, two or three fillers, and never by another experimental item.

3.2.1.3. Procedure

We monitored participants' eye movements using an Eyelink 1000 desktop head-stabilized tracker (SR Research). A 9-point calibration procedure was carried out at the beginning of the session, during the experiment additional recalibration was performed whenever necessary. Next, participants were presented 14 practice trials. Following the practice trials the experiment began. Each trial began with a black fixation dot in the middle of the screen (for drift correction). Once participants fixated this dot the experimenter initiated the trial. Figure 3.2 depicts the temporal order of an experimental trial.

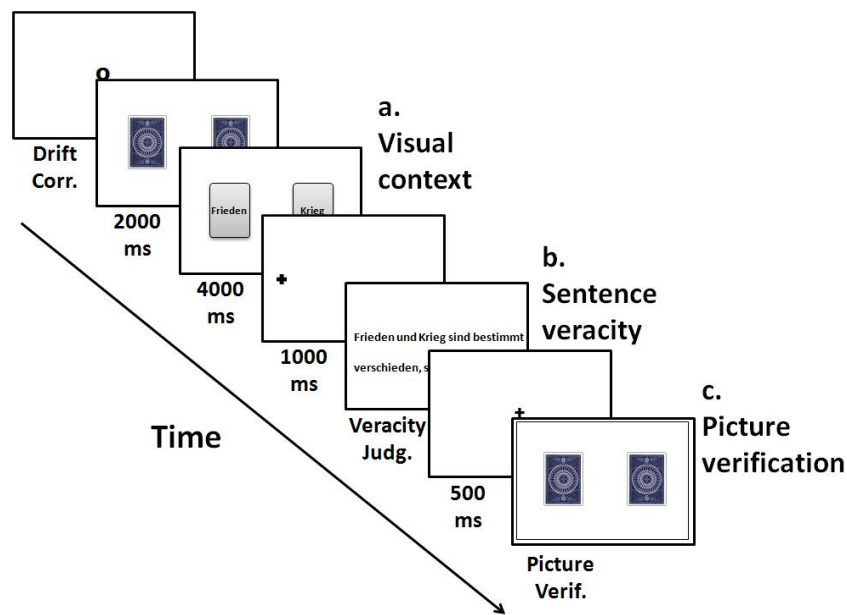


Figure 3.2. Schematic representation of an experimental trial for Experiment 1.

Every trial had three main steps: First, participants saw the backs of two playing cards that appeared from the bottom of the display moving upwards to the middle of the screen. In experimental trials, after the cards reached the middle of the screen they moved along the horizontal axis, either closer together or further apart, such that their final position was either close or far from one another (see Figure 3.1.). For all experimental trials and 24 filler trials the cards then turned and showed each an abstract written noun for 4000 ms (Fig. 3.2.a). For the remaining 72 fillers the cards turned and showed a blank front for 500 ms. Participants were instructed to look carefully at all of the cards and remember them. With regards to the word on the cards, participants were told that cards might sometimes present also written words on them. Next (Fig. 3.2.b), a black dot appeared for 1000 ms at the position where an ensuing sentence started. Participants were asked to focus the black dot, then carefully read the sentence and judge sentence veracity based on their world knowledge with a *yes* or a *no* button press (Cedrus Response Pad 8-Buttons, Large). In the third and final step (Fig. 3.2.c), participants verified whether a picture of two playing cards matched or mismatched the final position of the previously-inspected cards through a *yes* or a *no* button press.

3.2.1.4. Data analysis

Before statistical analyses, contiguous fixations below 80 ms duration were merged. Isolated fixations < 80 ms that could not be merged with other fixations or fixations >1200 ms were excluded (see, e.g., Sturt, Keller & Dubey, 2010). Trials that were incorrectly answered were also excluded from the analysis. Experimental sentences were divided into eight different areas of interest, as in (1). Our main analysis region was the adjective (ADJ) since it explicitly expresses the relation between the two words. We also examined the second noun phrase (NP2)

for possible earlier effects, and the second verb phrase (VP2) and the third noun phrase (NP3), to see whether predicted effects would manifest themselves in later sentence regions.

- (1) *Frieden*_{NP1} | *und*_{coord.} | *Krieg*_{NP2} | *sind*_{VP1} | *bestimmt*_{ADV} | *verschieden*_{ADJ} |
*das verriet*_{VP2} | *der Anthropologe*_{NP3}.
 ‘*Peace*_{NP1} | *and*_{coord.} | *war*_{NP2} | *are*_{VP1} | *certainly*_{ADV} | *different*_{ADJ} |
*suggested*_{VP2} | *the anthropologist*_{NP3}’

We computed first-pass time, regression path duration and total reading time. First-pass time is the sum of all fixations from first entering the region and prior to moving to another region. Regression path duration is the time elapsed from first entering a region until moving to the right of that region; unlike first-pass time it includes reading time following regressions out of the region (see, e.g., Konieczny, Hemforth, Scheepers & Strube, 1997; Liversedge, Paterson & Pickering, 1998; Rayner, 1998; Traxler, Pickering & Clifton, 1998). Finally, total reading time is the duration of all fixations in a given region (Rayner & Liversedge 2004; Rayner, 1998).

Visual inspection and the Kolmogoroff-Smirnov normality test confirmed that the data deviated from the normal distribution for all regions and conditions. To improve normality of the distribution, the raw data was log-transformed prior to inferential analysis (e.g., Tabachnik & Fidell, 2007, 246f.). Extreme data points for each interest region in each condition were further removed if their deletion improved normality (2.5% of the overall data and no more than 5% of the cases per condition). Log-transformation and removal of extreme cases improved skew and kurtosis, leading to non-reliable Kolmogorov-Smirnov statistics (KS $ps > .05$), with the exception of the NP2 region.

The data pertaining to log-transformed reading times was analysed using linear mixed-effect regression analyses (lme4 R). The benefits of mixed models in analysing psycholinguistic data are that these models allow simultaneous inclusion of participants and items as crossed random factors and an alternative to quasi- F , and separate by-participants and by-item analysis (F_1 , F_2 analyses). They also permit the inclusion of random slopes to model participant and item variation around the fixed effects. Furthermore, mixed models appear to be robust against missing values and do not require sphericity and homoscedasticity assumptions to be met (Baayen, Davidson & Bates 2008; Barr 2008, Quené & van den Bergh, 2004, 2008). The analysis produces estimates, standard errors and t -values for all main effects and their interaction. To evaluate significance we used the criteria of an absolute value of $t > 2$ (Baayen, 2008).

In our mixed model the two factors and their interactions (i.e., spatial distance and semantic similarity) constituted the fixed effects. To minimize collinearity, we centered the factors using a scale function⁶ in R. The scale function centers the data on a mean of 0 and a range of 2. Participants and items were included as crossed random intercepts. The omission of random slopes in linear mixed-effects regressions appears to be anti-conservative (Sturt et al., 2010) and risk increasing the possibility of Type I errors (Barr, Levy, Scheepers & Tily, 2013). Thus, we included all random slopes justifiable by our experimental design⁷. R-code for the full model: `lmer (dv ~ iv1*iv2 + (1+ iv1*iv2 | participant) + (1 + iv1*iv2 | item), data)`. In the R-code “dv” is the dependent variable (e.g., first-pass time), and “iv1” and “iv2” are the two independent variables (i.e., spatial distance and semantic similarity). This model includes fixed effects for the

⁶ `(data$factor=scale(as.numeric(data$factor))`

⁷ For measures in which we found interaction effects with the linear mixed-effect regression we also performed 2 x 2 repeated measures ANOVAs by participants and items. These support similar conclusions as the linear mixed-effects regression analyses (see Table B2 in Appendix B).

two factors, the interaction between them, participants and items as random intercepts, and both main and interaction effects as random slopes for both random intercepts.

Whenever the full model did not converge, we followed recommendations given by Barr and colleagues (2013), to increase the number of iterations in the estimation procedure, which resulted in convergence for all our models. We did not report pairwise comparisons since such analysis suffers from a considerable loss of power due to data splitting. Moreover, our within-subject within-item design is aimed to show interaction effects, if they exist. We also try to avoid increasing the family-wise error, due to repetitive analyses in a portion of the data.

3.2.2. Results

Table 3.2. shows mean reading times in milliseconds per condition, for all measures and regions analysed in Experiment 1. Results of the linear mixed-effect regression analyses are presented in Table 3.3.

Table 3.2. Mean reading times in milliseconds and standard errors of the mean (by condition, region and measure) in Experiment 1 (*continued on next page*).

Region	Condition	First-pass		Regression path		Total time	
		Mean	SE	Mean	SE	Mean	SE
NP2	Far-Dissimilar	174.23	3.5	220.48	8.3	436.54	23.2
	Close-Dissimilar	187.59	5.4	236.75	10.1	444.63	23.2
	Far-Similar	180.49	4.3	222.54	9.0	442.58	24.8
	Close-Similar	181.29	4.7	233.50	9.6	402.87	19.4
ADJ	Far-Dissimilar	328.32	7.7	408.58	12.5	416.30	10.9
	Close-Dissimilar	337.73	9.1	420.35	13.5	429.66	13.4
	Far-Similar	311.79	8.0	377.34	12.5	354.49	8.7
	Close-Similar	293.71	7.1	368.69	12.2	345.08	9.4
VP2	Far-Dissimilar	211.14	4.8	296.34	13.8	238.49	6.6
	Close-Dissimilar	239.54	7.1	307.95	12.7	289.99	14.0
	Far-Similar	218.60	5.7	312.94	14.1	253.19	8.8
	Close-Similar	212.28	5.6	340.24	16.8	250.95	8.0

NP3	Far-Dissimilar	333.13	14.1	973.70	59.7	404.82	17.3
	Close-Dissimilar	353.14	12.1	1006.58	61.3	418.43	15.3
	Far-Similar	340.58	16.2	964.22	57.9	394.51	18.9
	Close-Similar	318.26	11.1	849.97	45.4	367.87	13.4

Table 3.3. Main and interaction effects (by region and measure) in the linear mixed-effect regression on log-transformed reading times in Experiment 1.

Region	Fixed Effects	<i>First-pass</i>			<i>Regression path</i>			<i>Total times</i>		
		Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>
NP2	(Intercept)	5.130	0.024	211.76	5.274	0.037	143.06	5.739	0.068	84.18
	Distance	-0.011	0.011	-1.00	-0.022	0.017	-1.30	-0.011	0.025	-0.45
	Similarity	0.002	0.011	0.22	0.006	0.015	0.38	-0.002	0.032	-0.07
	Interaction	-0.010	0.012	-0.79	0.004	0.018	0.21	-0.015	0.020	-0.75
ADJ	(Intercept)	5.680	0.032	178.34	5.856	0.042	137.01	5.856	0.036	165.30
	Distance	0.009	0.013	0.74	0.000	0.013	0.01	0.009	0.012	0.74
	Similarity	0.047	0.018	2.63 *	0.058	0.018	3.38 *	0.085	0.019	4.46 *
	Interaction	-0.023	0.011	-2.08 *	-0.017	0.014	-1.18	-0.015	0.013	-1.18
VP2	(Intercept)	5.309	0.030	177.46	5.549	0.039	142.37	5.423	0.037	147.98
	Distance	-0.014	0.014	-0.96	-0.026	0.021	-1.21	-0.029	0.015	-1.95 #
	Similarity	0.021	0.013	1.78	-0.017	0.018	-0.93	0.017	0.015	1.09
	Interaction	-0.027	0.013	-2.09 *	0.001	0.022	0.04	-0.024	0.016	-1.56
NP3	(Intercept)	5.623	0.048	116.39	6.457	0.077	83.42	5.785	0.053	108.09
	Distance	-0.038	0.019	-1.96 #	-0.012	0.032	-0.38	-0.028	0.017	-1.62
	Similarity	0.008	0.017	0.47	0.012	0.028	0.43	0.020	0.021	0.98
	Interaction	-0.045	0.019	-2.39 *	-0.048	0.030	-1.61	-0.047	0.018	-2.58 *

Note: # $p < .1$. * $p < .05$.

At the critical adjective region (ADJ), we observed a pervasive similarity main effect across all reading time measures: reading times were shorter for sentences expressing similarity compared to those expressing dissimilarity (all t -values > 2). More importantly and as predicted, we found a reliable interaction between spatial distance and semantic similarity (t -value = -2.08) in first-pass reading time—that is, reading times were shorter when similarity-conveying sentences were preceded by cards-with-words presented close to each other (rather than far

apart) and reading times for sentences that conveyed dissimilarity were shorter when they were preceded by cards-with-words far apart (rather than close together). No other interaction effects were found in other measures in this region. We further analysed the second verb phrase (VP2) and the third noun phrase (NP3) regions. We found reliable interaction effects in both VP2 and NP3 regions in first-pass reading times (t -value = -2.09 and t -value = -2.39, respectively) and in addition, in total reading times for the NP3 region (t -value = -2.58), all of them with a similar pattern as the one found at the ADJ region in first-pass reading. Figure 3.3. illustrates the observed effect pattern between spatial distance and semantic similarity.

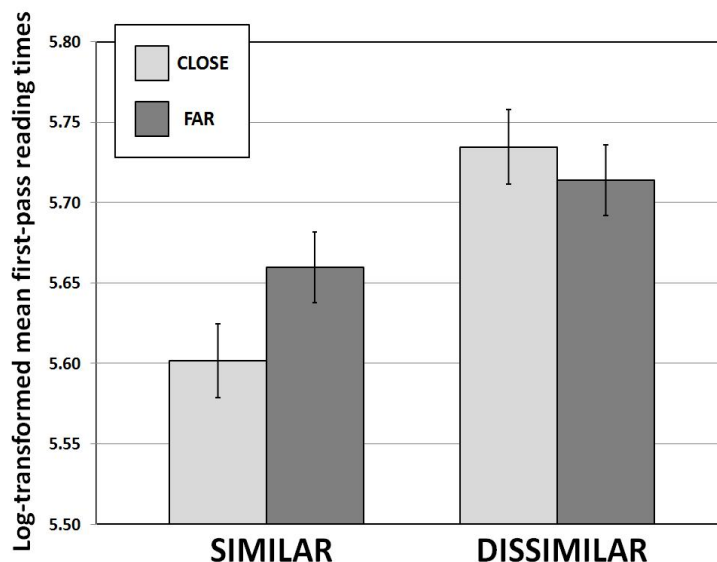


Figure 3.3: Log-transformed mean first-pass time (with error bars plotting the standard error of the mean) for the ADJ region as a function of sentence type and spatial distance between cards-with-words in Experiment 1.

3.2.3. Discussion

Spatial distance effects emerged at the ADJ and subsequent regions in first-pass times, suggesting they occur rapidly and incrementally. Effects at the VP2 and NP3 regions

corroborated the gaze pattern at the ADJ, and suggest spatial information can have an extended effect during sentence interpretation. Finally, the absence of reliable spatial distance effects before the adjective (at NP2) suggest spatial information can modulate semantic similarity between two abstract words as soon as it is explicitly mentioned (at the ADJ), but perhaps not when implied by coordinated noun phrases (e.g., '*peace and war...*')⁸. Results from Experiment 1 are to the best of our knowledge the first evidence showing that spatial distance can modulate incremental sentence interpretation (of similarity between two abstract nouns).

It is possible that rapid and incremental spatial distance effects occur when card distance is integrated with noun phrase similarity incrementally, during sentence comprehension. Indeed, previous findings have shown that semantic information from adjectives can be rapidly integrated with simultaneously presented non-linguistic visual information (e.g., Sedivy et al., 1999). Alternatively, the fact that the cards displayed words in Experiment 1 permitted pre-sentence integration of the words on the cards and card distance since the two abstract nouns on the cards implied either semantic similarity or dissimilarity; it is possible that this pre-sentence integration is responsible for the rapidity of spatial distance effects (in first-pass times) during sentence comprehension. Experiment 2 addressed this question.

3.3. Experiment 2

In Experiment 2 we separated the nouns from the spatial information conveyed by the visual context. If the rapid and incremental spatial distance effects depend upon pre-sentence

⁸ It is possible that the pre-sentence presentation of written nouns (which then re-appeared as NP1 and NP2 in the sentence) induced shallow semantic processing of the NP2. In fact, positive skew values for first-pass reading times at NP2 indicate short inspection times (skew after log-transform = .678, *SE*=.077). Repetition of words can speed up both word recognition and categorization (see Jacoby, 1983). This could be especially true for the NP2, since NP1 could have served as a cue of the previously seen word-pair. Such a reading strategy could have obscured potential spatial distance effects at the NP2 region on the assumption that deep semantic processing is pre-requisite for spatial distance effects on semantic interpretation.

integration of spatial distance of nouns, then we should observe no, or delayed, spatial distance effects when the nouns are presented temporally detached from the card context. Alternatively, if pre-sentence integration is not essential, then rapid and incremental spatial distance effects should emerge even when the nouns are presented temporally separated from the cards and the sentence trials.

We divided the experiment into six blocks. Before each block, participants learnt a list of written noun pairs (all critical pairs of that block and an equal number of filler word pairs). The cards that preceded each sentence, by contrast, remained blank. Participants thus saw the same noun pairs as in Experiment 1 but temporally removed from the cards and the sentences. As a result, they could neither predict the semantic relationships of a given sentence nor integrate the card distance with these semantic relations prior to sentence reading.

3.3.1. Method

3.3.1.1. Participants

Thirty-two further native German speakers with normal or corrected-to-normal vision (mean age: 23,19; range of age: 19 to 29) participated in the experiment for a monetary compensation of €6. None of them had been exposed to a second language before the age of 6. All participants gave informed consent.

3.3.1.2. Materials and design

Sentence stimuli and the design were identical to Experiment 1 but distinct to Experiment 1 cards on all of the critical trials were blank. We kept the experiment ratio of cards-with-words to blank cards constant by including written words on the cards of 72 filler trials.

Moreover, while blank cards were presented for 500 ms and cards-with-words for 4000 ms in Experiment 1, we presented both for 3000 ms in Experiment 2 to eliminate presentation-related differences between these trials.

3.3.1.3. Procedure and analysis

The analysis and procedure were identical to Experiment 1 with the following exception: Experiment 2 was divided in six blocks of 24 trials each. Before each block participants learnt a list of 16 word pairs (8 critical and 8 filler pairs). These pairs appeared in 16 of the 24 sentences in a block. Learning was self-paced and a memory test ensured a minimum of 75% accuracy in each list. The experiment lasted approximately one and a half hours.

3.3.2. Results

Accuracy on the test lists that participants had to perform before starting each of the six blocks was very high for all participants (mean = 96.8%, range = 95-100%). Table 3.4. shows mean reading times in milliseconds per condition, for all measures and regions analysed. At the NP2 region we found a main effect of similarity in regression path duration, but no other reliable effects were observed in this region. At the critical ADJ region we also observed a main effect of similarity for regression path duration, but we also found a marginal interaction effect between spatial distance and semantic similarity in regression path duration (t -value = -1.89) and a statistically reliable interaction effect in total reading times (t -value = -2.43). In these two measures, the ADJ region of sentences expressing similarity evidenced shorter reading times when preceded by cards that were presented close together compared to far apart, while the ADJ region of sentences that expressed dissimilarity showed shorter reading times when the previous

visual context presented cards far apart compared to close together. No effects were observed for the VP1 region. At the NP3 region, we only found a main effect of similarity for regression path duration and total reading times. No interaction effects were observed for these regions. Table 3.5. presents the results from the linear mixed effect regression for log-transformed reading time measures and sentence regions, both main and interaction effects of spatial distance and semantic similarity.

Table 3.4: Mean reading times in milliseconds and standard errors of the mean (by condition, region and measure) in Experiment 2.

Region	Condition	<i>First-pass</i>		<i>Regression path</i>		<i>Total time</i>	
		Mean	SE	Mean	SE	Mean	SE
NP2	Far-Dissimilar	229.46	5.18	298.05	11.34	465.67	19.51
	Close-Dissimilar	227.09	5.09	261.97	7.08	453.04	20.22
	Far-Similar	224.87	5.06	242.40	5.91	498.36	21.51
	Close-Similar	225.07	5.08	249.76	6.54	423.43	15.52
ADJ	Far-Dissimilar	330.47	7.48	402.16	10.98	389.54	10.02
	Close-Dissimilar	334.47	8.26	437.23	13.26	433.52	11.79
	Far-Similar	350.36	9.59	397.96	11.66	405.27	12.17
	Close-Similar	330.15	7.96	379.67	10.65	389.19	10.90
VP2	Far-Dissimilar	227.96	6.23	329.58	14.77	292.55	11.43
	Close-Dissimilar	227.15	5.53	300.67	11.15	283.03	9.56
	Far-Similar	240.62	6.73	289.34	9.15	311.67	12.70
	Close-Similar	227.65	5.56	297.63	11.08	318.00	13.75
NP3	Far-Dissimilar	354.40	12.29	951.21	48.98	426.59	16.44
	Close-Dissimilar	357.01	12.67	1028.25	57.45	432.37	16.24
	Far-Similar	380.99	14.66	1170.14	57.51	464.08	16.83
	Close-Similar	379.55	14.75	1152.59	64.42	466.12	17.79

3.3.3. Discussion

In Experiment 1 we found that spatial distance modulated semantic interpretation, and it did so rapidly and incrementally. Still, it was not clear whether these effects depended on

immediately preceding spatial-semantic integration. We detached the spatial information in the visual context from semantic information provided by the words, and replicated spatial distance effects in Experiment 2. As in Experiment 1, spatial distance effects did not emerge at NP2 but only later, once semantic similarity was explicitly mentioned (the ADJ region). Thus, spatial distance can rapidly affect the interpretation of semantic relations between abstract concepts when explicitly mentioned and even when integrated with sentence meaning as the sentence is being read.

Table 3.5: Main and interaction effect results (by region and measure) of the linear mixed-effect regression for log-transformed reading times in Experiment 2.

Region	Fixed Effects	<i>First-pass</i>			<i>Regression path</i>			<i>Total times</i>		
		Estimate	<i>SE</i>	<i>t</i>	Estimate	<i>SE</i>	<i>t</i>	Estimate	<i>SE</i>	<i>t</i>
NP2	(Intercept)	5.353	0.032	169.71	5.462	0.042	130.12	5.904	0.064	91.16
	Distance	0.003	0.014	0.22	0.007	0.015	0.43	0.025	0.016	1.57
	Similarity	0.011	0.011	0.96	0.042	0.018	2.39 *	-0.012	0.021	-0.57
	Interaction	-0.001	0.010	-0.12	0.019	0.012	1.49	-0.026	0.017	-1.54
ADJ	(Intercept)	5.730	0.031	184.08	5.888	0.034	173.67	5.895	0.038	155.86
	Distance	0.008	0.012	0.68	-0.009	0.012	-0.73	-0.017	0.012	-1.39
	Similarity	-0.004	0.016	-0.25	0.034	0.016	2.11 *	0.024	0.017	1.40
	Interaction	-0.015	0.013	-1.18	-0.027	0.014	-1.89 #	-0.034	0.014	-2.43 *
VP2	(Intercept)	5.357	0.031	170.54	5.561	0.040	140.78	5.530	0.045	121.80
	Distance	0.001	0.013	0.09	0.007	0.017	0.44	-0.009	0.016	-0.51
	Similarity	-0.009	0.013	-0.66	0.018	0.018	1.03	-0.022	0.021	-1.06
	Interaction	-0.010	0.012	-0.79	0.006	0.017	0.33	-0.000	0.017	0.02
NP3	(Intercept)	5.717	0.056	102.36	6.580	0.096	68.56	5.904	0.056	104.54
	Distance	-0.002	0.018	-0.1	0.003	0.023	0.14	-0.003	0.016	-0.21
	Similarity	-0.016	0.017	-0.92	-0.067	0.025	-2.73 *	-0.036	0.017	-2.18 *
	Interaction	-0.002	0.016	-0.13	-0.027	0.022	-1.26	-0.001	0.016	-0.04

Note: # $p < .1$. * $p < .05$.

While spatial distance effects emerged at the ADJ region in both experiments, they appeared earlier in Experiment 1 (first-pass reading times) than in Experiment 2 (marginal in regression path duration, and significant in total reading times). This delay could have resulted from the immediate integration of card distance with the semantic relationships of the noun phrases. Alternatively (or in addition), the cards and/or the first noun phrase of the sentence may have cued retrieval of previously seen or learnt semantic relations. This retrieval process and integration of learnt noun pairs with the sentence content and spatial distance information could have delayed the emergence of spatial distance effects.

3.4. Experiment 3

To exclude this, the cards for all trials were blank in Experiment 3, and participants saw no word pairs prior to sentence reading. Thus, upon encountering the cards or even the first noun phrase, participants had no information about the semantic relationship expressed by the sentence. If retrieval processes delayed spatial distance effects in Experiment 2 but otherwise the integration of spatial distance and semantic relations during comprehension is rapid, then we should observe spatial distance effects in first-pass times at the ADJ in Experiment 3.

It is worth noting that Experiment 3 examined spatial distance effects on semantic interpretation in the absence of any links (e.g., referential or lexical associations) other than temporal contiguity (the cards appeared immediately before the sentence). Furthermore, cards were irrelevant to the sentence comprehension task in all three experiments. Experiment 3 thus provides a strong test of whether spatial distance alone can modulate the incremental semantic interpretation of similarity relations.

3.4.1. Method

3.4.1.1. Participants

Another thirty-two native German speakers with normal or corrected-to-normal vision (mean age: 24,37; range of age: 20 to 31) participated in the experiment for a payment of €6. None of them had been exposed to a second language before the age of 6. All participants gave informed consent.

3.4.1.2. Material, design, procedure and data analysis.

The sentence stimuli, experimental design, procedure and data analysis were the same as in Experiment 1 but cards on all trials were blank in Experiment 3. As in Experiment 2, cards turned around after 3000 ms and presented the front for another 3000 ms.

3.4.2. Results

Table 3.6. shows mean reading times in milliseconds per condition, for all measures and regions analysed. Table 3.7. shows the results from the linear mixed effect regression on log-transformed reading times. At the NP2 we found a marginal main effect of similarity in first-pass reading time (t -value = 1.94) and, in addition, we found a significant interaction effect between spatial distance and semantic similarity (t -value = -2.15) in this measure. First-pass reading time for sentences that expressed similarity were shorter when they were preceded by cards close to each other (compared to far apart), while first-pass reading time for sentences that expressed dissimilarity were shorter when preceded by cards far apart (compared to close together). No other effects were observed for this region. At the critical ADJ region we observed a statistically significant main effect of similarity for regression path duration. Yet, no interaction effects were

observed in this region. Analysis of the VP2 showed a significant interaction effects between spatial distance and semantic similarity (t -value = -2.07), but no other effects were found in this region. Finally, the NP3 region evidenced main effects for both spatial distance and semantic similarity in regression path duration, but no interaction effects were found in this region. Both the reliable interaction effect at the NP2 and at the VP2 regions exhibit a similar pattern of interaction as the one reported in Experiment 1.

Table 3.6: Mean reading times in milliseconds and standard errors of the mean (by condition, region and measure) in Experiment 3.

Region	Condition	<i>First-pass</i>		<i>Regression path</i>		<i>Total time</i>	
		Mean	SE	Mean	SE	Mean	SE
NP2	Far-Dissimilar	313.71	9.57	434.60	19.91	614.25	23.43
	Close-Dissimilar	354.92	11.78	411.96	14.49	614.59	22.63
	Far-Similar	300.14	8.42	415.25	19.59	608.56	27.23
	Close-Similar	290.28	8.63	382.56	19.37	635.67	29.73
ADJ	Far-Dissimilar	363.66	8.99	436.77	12.19	427.85	11.24
	Close-Dissimilar	350.02	9.72	428.57	13.02	442.61	14.23
	Far-Similar	334.79	9.08	389.74	11.66	409.38	11.91
	Close-Similar	343.45	9.96	405.83	13.65	432.26	14.87
VP2	Far-Dissimilar	229.47	5.86	297.56	10.44	269.98	8.42
	Close-Dissimilar	249.01	7.15	334.17	13.59	308.65	11.09
	Far-Similar	239.95	6.16	336.13	13.38	299.32	10.68
	Close-Similar	236.70	6.80	337.50	14.88	331.05	16.33
NP3	Far-Dissimilar	393.35	15.44	1095.71	61.03	475.64	19.26
	Close-Dissimilar	419.48	17.99	1131.12	57.81	497.81	19.62
	Far-Similar	402.99	15.75	1198.33	63.58	526.88	20.49
	Close-Similar	411.50	18.62	1283.04	70.90	514.92	20.56

3.4.3. Discussion

Analyses of the data from Experiment 3 revealed rapid (first-pass) and incremental (at the NP2 region) interaction effects between spatial distance and semantic similarity. These findings provide evidence that even in the absence of any overt relation between a visual context and the

content of the subsequent sentence, non-linguistic spatial information can affect the semantic interpretation of conjoined abstract nouns.

Table 3.7: Main and interaction effects (by region and measure) of the linear mixed-effect regression for log-transformed reading times in Experiment 3.

Region	Fixed Effects	<i>First-pass</i>			<i>Regression path</i>			<i>Total times</i>		
		Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>
NP2	(Intercept)	5.629	0.050	112.38	5.820	0.056	104.74	6.195	0.068	91.32
	distance	-0.017	0.015	-1.16	0.012	0.018	0.68	-0.024	0.020	-1.18
	similarity	0.036	0.019	1.94 #	0.023	0.022	1.07	-0.002	0.024	-0.09
	interaction	-0.029	0.013	-2.15 *	-0.018	0.021	-0.84	0.000	0.017	0.02
ADJ	(Intercept)	5.752	0.038	153.12	5.917	0.037	159.68	5.935	0.045	131.99
	distance	0.011	0.012	0.91	0.005	0.013	0.38	-0.003	0.013	-0.25
	similarity	0.026	0.018	1.43	0.050	0.021	2.38 *	0.024	0.017	1.38
	interaction	0.019	0.012	1.56	0.014	0.014	0.99	0.003	0.013	0.24
VP2	(Intercept)	5.381	0.034	160.22	5.624	0.041	136.85	5.546	0.046	120.43
	distance	-0.004	0.016	-0.27	-0.005	0.020	-0.28	-0.022	0.017	-1.31
	similarity	-0.001	0.012	-0.07	-0.020	0.017	-1.21	-0.017	0.017	-1.00
	interaction	-0.025	0.012	-2.07 *	-0.031	0.020	-1.51	-0.026	0.018	-1.46
NP3	(Intercept)	5.776	0.056	103.23	6.691	0.084	79.56	5.975	0.069	86.59
	distance	-0.025	0.019	-1.31	-0.056	0.024	-2.31 *	-0.026	0.017	-1.50
	similarity	0.007	0.019	0.39	-0.058	0.024	-2.41 *	-0.030	0.019	-1.57
	interaction	-0.009	0.019	-0.53	0.001	0.025	0.05	-0.014	0.018	-0.75

Note: # $p < .1$. * $p < .05$.

Spatial distance effects appeared at the NP2 region in Experiment 3, therefore earlier in the sentence than in the first two experiments. It is possible that the absence of interaction effects at NP2 in the first two experiments resulted from the repetition of the words. Seeing the noun phrases prior to the sentence might have prompted comprehenders to process them more superficially during sentence reading, such that their meaning was not immediately integrated with spatial distance information. Furthermore, the earlier effects in Experiment 3 (at NP2) than Experiment 2 (in regression path times at the ADJ region) suggest that the relative delay in

Experiment 2 was unlikely caused by immediate integration of spatial distance and semantic similarity and rather resulted from cued retrieval of previously-learned semantic relations.

One question remains open; it is unclear why spatial distance effects were present at the ADJ region in Experiments 1 and 2 but absent from that region in Experiment 3. One possibility is that integration of spatial information and implicit semantic similarity at NP2 has driven attention to spatial distance in such a way that it eliminated spatial distance effects at the downstream ADJ region. Tentative evidence for the view that spatial distance affected semantic similarity processing differently at the ADJ region in Experiment 3 comes from descriptively different gaze pattern at that region (first-pass reading times and regression path duration for similarity-conveying sentences were longer when cards were close together than far apart, and vice versa for sentence expressing dissimilarity). Only after processing the adjective that explicitly showed the two abstract nouns to be similar or different (at the VP2 region), the effect of spatial distance on semantic interpretation re-appears as in Experiment 1.

3.5. Summary of results

In three visually situated reading experiments we addressed a question with regards to the potential influence of spatial distance on semantic similarity interpretation, beyond referential and lexical-semantic links between a sentence and a visual context. In Experiment 1, we observed a distinctive effect of spatial distance (close together vs. far apart) on early reading measures, from the moment in which semantic similarity is made explicit in the sentence (at the ADJ region), and extended in subsequent regions (i.e., VP2 and NP3 regions): first-pass reading times were shorter for sentence expressing similarity when preceded by cards close together, compared to when cards were far apart. Instead, for sentence expressing dissimilarity shorter

reading times were observed when cards presented before the sentence were far apart, compared to when cards were presented close together.

In Experiment 2, we replicated this pattern of effect, but in later measures (regression path duration), and the effect was restricted to the adjective region. In Experiment 3, we again observed the same interaction pattern as in previous two experiments, however, this time in an earlier in the sentence region. Even before semantic similarity is made explicit in the sentence, spatial distance between objects modulated first-pass reading times at the NP2 region, where semantic similarity is implied by the and-coordination of abstract nouns. This effect re-appeared after the semantic similarity is made explicit (at the ADJ region), in the VP2 region and in first-pass reading times. Overall, our results are the first evidence of rapid and incremental effects of spatial distance on the semantic interpretation of abstract language.

reading times were observed when cards presented before the sentence were far apart, compared to when cards were presented close together.

In Experiment 2, we replicated this pattern of effect, but in later measures (regression path duration), and the effect was restricted to the adjective region. In Experiment 3, we again observed the same interaction pattern as in previous two experiments, however, this time in an earlier in the sentence region. Even before semantic similarity is made explicit in the sentence, spatial distance between objects modulated first-pass reading times at the NP2 region, where semantic similarity is implied by the and-coordination of abstract nouns. This effect re-appeared after the semantic similarity is made explicit (at the ADJ region), in the VP2 region and in first-pass reading times. Overall, our results are the first evidence of rapid and incremental effects of spatial distance on the semantic interpretation of abstract language.

CHAPTER 4

Can spatial distance modulate semantic interpretation of social relations incrementally?

4.1. Introduction

In Chapter 3 we examined whether semantic interpretation of abstract sentences could be rapidly and incrementally modulated by non-linguistic information, and if non-linguistic visual information could modulate language comprehension even in the absence of referential or lexical-semantic links. In Chapter 4 we address two further questions with regards to the relationship between spatial information and abstract language. Specifically we investigated if (a) spatial distance could modulate semantic interpretation about social relations and (b) whether effects of spatial distance on semantic interpretation could be observed even if two objects are presented in the visual context, but when no coordination structure appears in the subsequent sentence. Finding rapid and incremental interactions between spatial distance and social relations, would suggest that abstract semantic domains, other than similarity, can interact rapidly and incrementally with spatial distance. Additionally, the time course of these potential

effects will shed light on the co-indexing process through which abstract language interacts with spatial information.

4.1.1. The present study

The review presented in Chapter 2 shows that connection between social relations and spatial distance (defined as INTIMACY IS CLOSENESS by conceptual metaphor) appears as a good candidate to further examine the role of visual information in online abstract sentence interpretation. According to conceptual metaphor, expressions such as “Your research topic is *close* to mine” are metaphorical in the sense that they use a concrete spatial adjective (i.e., *close*) to refer to an abstract idea, namely “Your research topic is *similar* to mine”. When we examined if this metaphorical mapping between spatial distance and similarity could be observed during online sentence processing, we found that such effects occurred rapidly and incrementally (in first-pass reading times and as the sentence unfolded). In Experiment 1, we first observed the predicted spatial distance effects at the adjective region, when similarity (or dissimilarity) was first explicitly mentioned. In Experiment 3, we observed such effects at the second noun region, suggesting that space can modulate similarity even before being explicitly mentioned, but rather when it was implied by the synonyms (or antonyms) in noun *and*-coordinations. Our findings extended previous studies on the connection between spatial distance and similarity, from rating and response times to incremental language comprehension (cf. Boot & Pecher, 2010; Casasanto, 2008), and they replicate the time course (i.e., first-pass reading time) of previous findings on the effects of visual features on reading times for concrete language (see Wassenburg & Zwaan, 2010). They are also congruent with the linguistic focus hypothesis (Taylor & Zwaan, 2008) and the co-indexing mechanism between language and visual representations (Knoeferle & Crocker,

2007; see also Glenberg & Robertson, 1999). However, while co-indexing for a concrete word (e.g., “apple”), with its visual reference (a picture of an apple) is fairly straight forward, co-indexing between abstract words and spatial information is much less clear.

The sentence structure on our first set of experiments presented a noun phrase *and*-coordination, followed by a verb, an adverb and a critical adjective, as in ‘*Peace*_{NP1} *and*_{coord.} *war*_{NP2} *are*_{VP1} *certainly*_{ADV} *different*_{ADJ...}’. There are at least two possibilities with regards to the co-indexing between spatial information and semantic similarity. One possibility was that each of the two sentential noun phrases could co-index with each of the two playing cards presented in the visual context (and not necessarily spatial distance *per se*). This is particularly plausible in Experiment 1, when cards turned around and showed the two sentential nouns in the visual context, but it is also plausible even when no words were presented in the visual context, as the number of objects matched the number of nouns in the *and*-coordination. Alternatively, the abstract semantic representation of similarity, implied by the noun phrase *and*-coordination and made explicit by the adjective, could be co-indexed with the spatial information (distance between objects), without it being necessary that each object in the visual context is co-indexed with each of the nouns in the sentence. Experiment 4 addressed this question.

4.2. Experiment 4

The data presented in Chapter 3 could not reveal if there was co-indexing between the two playing cards and the two abstract words or if, it was due to perceived distance and the understanding of similarity. To address this issue in Experiment 4, instead of the *and*-coordinated nouns in the subject of the sentence (Experiments 1, 2 and 3), the critical sentence

had a single noun as subject, followed by a verb, an adverb, and another noun as object of the verb (see Table 4.1.). We predicted that if co-indexing of abstract language and visual information requires a match between the objects in the visual context (two playing cards) and the nouns of the sentence, we should see the effect of spatial distance only after the second noun phrase of the sentence has been visited. However, if abstract language and visual information co-index, even if the number of objects in the visual context does not match the number of nouns in the subject of the sentence, then we should see early (first-pass or regression path duration) effects of spatial distance at the adverb region.

4.2.1. Method

4.2.1.1. Participants.

Thirty-two native German speakers (mean age: 23,93; range of age: 19 to 31) with normal or corrected-to-normal vision from the Bielefeld University community each received €6 Euros for participating. None of them had been exposed to a second language before age 6. All participants gave informed consent.

4.2.1.2. Materials and design.

The visual display and the design were identical as in Experiment 1. Two playing cards appeared on the screen and moved, in experimental trials, close together (the two cards moved from their initial middle position closer to each other) or far apart (the two cards moved from their initial position further apart from each other). After reaching their final position cards were turned over and revealed the two first sentential nouns of an ensuing sentence. However, we

constructed 48 new critical sentences expressing interactions between two characters. Each of the sentences had two versions, one expressing a friendly encounter (*‘Sandra met cheerfully her aunt at the health center’*) or an unfriendly encounter (*‘Sandra met grumpily her aunt at the health center’*). Only the modal adverb (e.g., “cheerfully” vs. “grumpily”) was different in the two sentences, which were matched in frequency and length. The combination of the factors spatial distance (Close or Far) and social relation (Friendly or Unfriendly) resulted in four experimental conditions. Table 4.1. summarizes the experimental condition, and gives an example of the critical sentences and regions of interest.

Table 4.1. Example item sentences and the four conditions

Cards	Sentence	Condition
FAR	Sandra _{NP1} trifft _{VP} gutgelaunt _{ADV} ihre _{pron.} Tante _{NP2} in der Praxis _{pp.}	Far-Friendly
CLOSE	Sandra _{NP1} trifft _{VP} gutgelaunt _{ADV} ihre _{pron.} Tante _{NP2} in der Praxis _{pp.}	Close-Friendly
FAR	Sandra _{NP1} trifft _{VP} missmutig _{ADV} ihre _{pron.} Tante _{NP2} in der Praxis _{pp.}	Far-Unfriendly
CLOSE	Sandra _{NP1} trifft _{VP} missmutig _{ADV} ihre _{pron.} Tante _{NP2} in der Praxis _{pp.}	Close-Unfriendly

Translation. Friendly sentence condition: ‘Sandra met cheerfully her aunt at the health center’;
Unfriendly sentence condition: ‘Sandra met grumpily her aunt at the health center’

Each participant was presented with the same number of trials per condition and with every experimental item in only one condition (Latin square design). A number of 96 filler sentences were also presented in each experiment. In total, participants took part in 144 trials, out of which half presented the two first sentential nouns (e.g., “Sandra”; “Tante”) on the front of the cards (24 of the fillers, and all critical trial). The other half presented a blank front. The visual context of fillers and the trial presentation order emulated Experiment 1.

4.2.1.3. Procedure.

The procedure was identical to Experiment 1.

4.2.1.4. Data analysis

The data analysis was almost identical to Experiment 1, except for the regions of interest and the outlier removal threshold. In Experiment 4, experimental sentences were divided into six different areas of interest, as in (2). Our main analysis regions were the adverb (ADV), since it explicitly expresses the manner of the social interaction, and the second noun phrase (NP2), in order to assess our research question regarding co-indexing. Both are marked in bold.

(2) “Sandra_{NP1} | trifft_{VP} | **gutgelaunt**_{ADV} | ihre_{pron.} | **Tante**_{NP2} | in der Praxis_{SP}”.

*‘Sandra_{NP1} | met_{VP} | **cheerfully**_{ADV} | her_{pron.} | **aunt**_{NP2} | at the health center_{PP}.;*

First-pass time, regression path duration and total reading time were computed using Data Viewer software (SR Research). Reading times longer or shorter than 3 SD from the mean per participant were removed (less than 1.3% of the data). Further, the raw data was log-transformed prior to inferential analysis. Outlier removal and log-transformation improved the residuals distribution in the subsequent linear regression, leading to non-reliable Kolmogorov-Smirnov statistics (KS ps > .05)⁹.

⁹ The removal of a larger proportion of outliers (e.g., 2.5 SD or 2 SD) proved to have the opposite effect on the distribution of the residuals, namely the distribution moved further away from normal distribution.

4.2.2. Results

Table 4.2. shows mean reading times in milliseconds per condition, for all three measures in the two critical regions. At the ADV region we observed no main effects of distance or similarity in any reading measure. However, a reliable interaction between spatial distance and the social relation ($t = 2.03$) was found in first-pass reading time: reading times were shorter for sentences expressing unfriendly relations when preceded by cards close together (compared to far apart), while reading times for sentences expressing friendly relations were shorter when preceded by cards far apart (compared to close together). No other significant interactions were found in this sentence region. At the NP2 region a main effect of social relation in total reading times was observed; sentences that expressed unfriendly relations revealed longer reading times compared to those sentences expressing friendly relations ($t = 2.27$). A summary of the results from the linear mixed-effect regression analyses is presented in Table 4.3. Figure 4.1. illustrates the interaction effect between spatial distance and semantic similarity.

Table 4.2. Mean reading times in milliseconds and standard errors of the mean (by condition, region and measure) in Experiment 4.

Region	Condition	<i>First-pass</i>		<i>Regression path</i>		<i>Total time</i>	
		Mean	SE	Mean	SE	Mean	SE
ADV	Far-Unfriendly	289.04	7.9	353.81	13.1	488.86	19.6
	Close-Unfriendly	266.73	7.1	353.96	13.6	464.12	18.5
	Far-Friendly	278.80	6.9	351.31	11.2	427.90	15.3
	Close-Friendly	299.31	8.3	363.46	11.4	461.68	16.8
NP2	Far-Unfriendly	251.93	9.0	320.40	14.2	404.14	21.2
	Close-Unfriendly	235.18	6.9	314.85	14.4	426.09	27.1
	Far-Friendly	234.17	7.9	289.54	12.3	382.57	19.2
	Close-Friendly	219.83	6.4	308.33	19.0	368.48	23.3

Table 4.3. Main and interaction effects (by region and measure) in the linear mixed-effect regression on log-transformed reading times in Experiment 4.

Region	Fixed Effects	<i>First-pass</i>			<i>Regression path</i>			<i>Total times</i>		
		Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>
ADV	(Intercept)	5.567	0.039	143.66	5.744	0.042	136.17	5.968	0.054	110.01
	distance	0.002	0.013	0.19	-0.006	0.017	-0.33	-0.007	0.016	-0.47
	similarity	-0.014	0.013	-1.02	-0.010	0.016	-0.61	0.033	0.024	1.38
	interaction	0.027	0.013	2.03 *	0.011	0.015	0.74	0.026	0.017	1.55
NP2	(Intercept)	5.345	0.040	134.62	5.533	0.051	108.74	5.698	0.071	80.5
	distance	0.020	0.015	1.30	-0.003	0.022	-0.15	0.007	0.019	0.38
	similarity	0.028	0.015	1.82 #	0.026	0.022	1.18	0.051	0.022	2.27 *
	interaction	0.002	0.014	0.17	0.006	0.018	0.33	-0.014	0.018	-0.80

Note: #*p* < .1. **p* < .05.

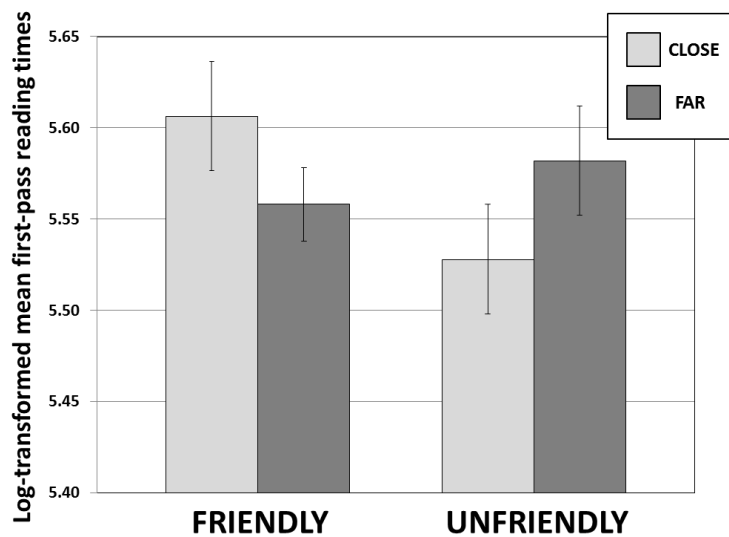


Figure 4.1.: Log-transformed mean first-pass time (with error bars plotting the standard error of the mean) for the ADV region as a function of sentence type and spatial distance between cards-with-words in Experiment 4.

4.2.3. Discussion

As with results from Experiments 1 and 3, distinctive spatial distance effects on semantic interpretation emerged rapidly (in first-pass reading times) and incrementally (as the sentence unfolded). These results suggest that perceived spatial distance can modulate real-time sentence interpretation as a function of the kind of social relation that the sentence describes. Furthermore, the observed interaction effect appears first at the ADV region, and crucially in first-pass reading times, which suggests that spatial distance between two objects can co-index abstract language even without an *and*-coordination, and without the need of mapping subject nouns to objects in the visual context. However, two clear differences appear when comparing the results from Experiment 4 with previous experiments described in Chapter 3. First, we found a focal effect, compared to the extended effect in first-pass reading times found in Experiment 1 (no interactions were observed after the first-pass at the ADV region; cf. Chapter 3, Experiment 1). This could be interpreted as having a weaker co-index without a match between the number of objects and the nouns at the sentence subject. Otherwise (or additionally), it could be related to a different degree of mapping between spatial distance and the semantic domain of social relations compared to similarity.

The second main difference is the pattern of these effects. In contrast to experiments presented in Chapter 3 (and the behavioral data discussed in Chapter 2), the result pattern in Experiment 4 showed an advantage in processing incongruent conditions (i.e., close-unfriendly and far-friendly) and a disadvantage for congruent conditions (i.e., close-friendly and far-unfriendly), which we will refer to as *interference* effect. Interference (vs. facilitation) interaction between conceptual and perceptual processing has been previously reported in the literature. For example, Richardson et al. (2003) reported two experiments that investigated how

the axis activated by a sentential verb (see Richardson, Spivey, Edelman & Naples 2001) affect a subsequent perceptual task. In Experiment 1, participants heard sentences after which they were presented with a neutral visual cue (a circle or a square). Participants had to discriminate whether the cue was a circle or a square. Results showed that response latencies were longer when the visual cue was presented in the same axis (vs. different) as that elicited by verb of the sentence (e.g., “The car impacts the wall”: horizontal axis; “The plane bombs the city”: vertical axis), which was described as an interference effect.

In Experiment 2, participants listened to a set of sentences presented simultaneously with a sequence of pictures depicting the subject and object of the each sentence. In a subsequent picture-memory task, participants were faster to recognize the pictures arranged horizontally when the sentential verb implied horizontality (vs. verticality), and faster for pictures arranged vertically when the sentential verb implied verticality (vs. horizontality). Contrary to Experiment 1, participants responded faster in a picture-memory task, when the sentential verb orientation matched the presentation of the pictures. Thus this effect was described as facilitation.

Existing accounts of interference (and facilitation) interaction between conceptual and perceptual processing (see Bergen, 2007; Connell & Lynott, 2012; Kaschak et al., 2005; Richardson et al., 2003) argue for a variety of moderating factors. Among them, the role of presentation timing (sequential vs. simultaneous), integrability, and attentional modulation have been discussed. According to Kaschak and colleagues (2005), interference effects should appear when both conceptual and perceptual information are presented simultaneously but they are non-integrable. The authors’s concept of integrability generally refers to whether the linguistic content directly refers to objects visually depicted. For instance, “the sentence and the [visual] stimulus are integrable [...] when one sees the image of a car while processing the sentence,

“The car approached you”.” (Kaschak et al., 2005: p. B86). By contrast, seeing an animation of a black-and-white spiral, creating an illusion of forward or backward movement, is not integrable with the content of a sentence, “The car approached you”, since there are no concrete objects to refer to in the visual depiction.

However, Bergen et al. (2007) and Richardson et al. (2003), reported interference effects in an experiment with the sequential presentation of stimuli (see Richardson et al., Experiment 1 described above), and when stimuli were non-integrable. If interference effects can occur both with sequential or simultaneous presentation, the interaction effect observed in Experiment 4 could have emerged from the non-integrability of spatial distance and the social relation expressed in the sentences. However, to accommodate the effects observed in Experiments 1, 2 and 3, this account would have to assume that semantic similarity is integrable with spatial distance, but social relations are not. This is potentially true, if in contrast to semantic similarity, social relation have a wider range of experiential correlates besides spatial distance. In fact, while semantic similarity is almost exclusively mapped into concrete concepts about distance, social relation can be described using a number of other concrete domains. For instance, hierarchical relations are commonly conceptualized in terms of verticality (e.g., “He ranks above me”, “They are lower class”), and affective relations can be expressed in terms of temperature (e.g., “They gave me a warm welcome”; see Lakoff et al., 1991).

A more recent account focused on attentional modulation rather than integrability and timing of presentation (Connell & Lynott, 2012). This alternative account could also accommodate our findings. This account explicitly assumes that both facilitation and interference can occur with a simultaneous or sequential presentation of visual and linguistics stimuli, and when stimuli are apparently integrable as well as when they are not integrable. According to

Connell & Lynott, interference effects appear, for instance, if perceptual stimuli (e.g., spatial distance between cards) occupy modality-shared attentional resources needed to process similar conceptual information (e.g., social relations). This means that as long as the focus of attention remains on the representation of, for instance spatial closeness, interpretation of a related concept (e.g, friendliness), would be slower compared to an unrelated concept (e.g., unfriendliness). Facilitation, instead, would be observed when stimuli initially direct attention to a perceptual modality, for instance spatial closeness, and soon after that the focus of attention moves from that perceptual representation to the linguistic content, resulting in priming for related concepts such as friendliness, compared to unfriendliness.

A potential interpretation, in this case, is that in Experiment 4 card distance occupies attentional resources extensively, or in other words the focus of attention remained with the perceptual representation, while in Experiments 1, 2 and 3, the focus of attention rapidly drifts from the spatial distance to semantic interpretation. Indeed, in the first three experiments, the number of nouns and the number of cards matched before the critical adjective appeared in the sentence. By contrast, in the present study one of the cards was related (through a written word) to the sentence subject and the second card was related to the sentence object. In this context, it is plausible that when participants have to process the critical adverbial region, after the word on the first card appears but before the word on the second card appears, participants' attention could be directed extensively to the spatial distance between cards leaving fewer resources to process the associated adverb. Experiment 5 is aimed to disentangle these two models, and to investigate the effects of non-referential (i.e., fully unrelated) visual context on abstract sentence interpretation further.

4.3. Experiment 5

Although both of the accounts described above focused on different aspects of the relationship between the perceptual and the conceptual processing, both can accommodate the results from Experiment 4. An integrability account concentrates on the integration between perceptual and semantic content, namely whether they are integrable (a picture of an apple and the word “apple”), or non-integrable (a circle and the word “apple”, for instance). An attentional account instead concentrates on the representation that is more active at a certain point, namely where does the focus of attention rest as the semantic content of the sentence is being processed. Thus, the two accounts would make different predictions if the sentence content does not change but when some attentional cues in the visual context are altered.

In Experiment 5 we address this issue by keeping the same content (same sentences), but removing all words from the cards in the visual context. As discussed above, attention was potentially driven to the mental representation of spatial distance between cards by the word on the first card. Thus, presenting no words on cards would remove the critical attentional cue in the visual context, such as attention should remain focused on semantic interpretation. In this context, we predict a facilitation effect of spatial distance on reading times. On the other hand, if the interference effect arises from the non-integrability of social relations and card distance, we should again observe an interference-like effect as in Experiment 4. In addition, by removing all words from the cards in the visual context, we can further test the hypothesis put forward in Chapter 3: non-referential visual context can modulate real-time sentence interpretation.

4.3.1. Method

4.3.1.1. Participants.

Another thirty-two native German speakers with normal or corrected-to-normal vision (mean age: 25,12; range of age: 19 to 34) participated in the experiment for a monetary compensation of €6 Euros. None of them had been exposed to a second language before age 6. All participants gave informed consent.

4.3.1.2. Material, design, procedure and data analysis.

The sentence stimuli, experimental design, procedure and data analysis were the same as in Experiment 4 but cards on all trials were blank in Experiment 5.

4.3.2. Results

Table 4.4. shows mean reading times in milliseconds per condition, for all three measures in the two critical regions, and a summary of the results from the linear mixed-effect regression analyses is presented in Table 4.5. At the ADV region no main effects of distance or similarity, nor interaction effects between distance and similarity were observed in any reading measure. As can be seen in Table 5, regression path duration evidence a facilitation-like pattern; longer reading times were observed for sentences expressing unfriendly relations when preceded by cards far apart compared to close together, while longer reading times were observed for sentences expressing friendly relations when preceded by cards close together (compared to far apart). However, this effect pattern was not statistically significant. Similarly, no effects of distance, similarity or their interaction were observed at the NP2 region.

Table 4.4. Mean reading times in milliseconds and standard errors of the mean (by condition, region and measure) in Experiment 5.

Region	Condition	<i>First-pass</i>		<i>Regression path</i>		<i>Total time</i>	
		Mean	SE	Mean	SE	Mean	SE
ADV	Far-Unfriendly	305.62	8.5	351.68	10.7	506.30	16.6
	Close-Unfriendly	303.86	9.1	361.83	11.9	525.75	21.6
	Far-Friendly	312.42	7.7	354.14	10.8	477.45	16.8
	Close-Friendly	296.89	8.0	343.37	10.8	502.34	18.4
NP2	Far-Unfriendly	296.69	10.7	360.57	14.3	474.14	20.8
	Close-Unfriendly	265.42	8.2	354.70	18.4	443.82	19.4
	Far-Friendly	282.13	9.4	367.96	17.2	445.85	20.0
	Close-Friendly	275.30	7.8	359.31	14.8	433.41	19.8

Table 4.5. Main and interaction effects (by region and measure) in the linear mixed-effect regression on log-transformed reading times in Experiment 5.

Region	Fixed Effects	<i>First-pass</i>			<i>Regression path</i>			<i>Total times</i>		
		Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>	Estimate	SE	<i>t</i>
ADV	(Intercept)	5.614	0.041	138.41	5.737	0.043	131.99	6.038	0.053	113.58
	distance	0.011	0.012	0.92	0.003	0.014	0.2	0.021	0.018	1.17
	similarity	-0.007	0.012	-0.58	0.009	0.015	0.61	-0.007	0.016	-0.4
	interaction	-0.012	0.014	-0.89	-0.014	0.015	-0.92	0.016	0.019	0.81
NP2	(Intercept)	5.499	0.043	129.05	5.691	0.057	100.49	5.873	0.066	88.4
	distance	0.009	0.015	0.57	-0.002	0.017	-0.13	0.03	0.019	1.56
	similarity	-0.003	0.014	-0.25	-0.001	0.017	-0.08	0.012	0.017	0.73
	interaction	0.019	0.014	1.38	0.009	0.016	0.54	0.011	0.019	0.58

Note: #*p* < .1. **p* < .05.

4.3.3. Discussion

Experiment 5 aimed to contrast the predictions of two accounts of interference and facilitation interactions between perceptual and conceptual processing. According to our interpretation of such accounts, both can accommodate the interference effects observed in

Experiment 4, however, they would make different predictions for Experiment 5: an integrability-based account (Kaschak et al., 2005; Bergen, 2007) would predict interference-like interaction effects, while an attentional-based account (Connell & Lynott, 2012) would predict a facilitation-like interaction effect if words are removed from the cards in the visual context. None of these accounts would predict null effects under these conditions. Additionally, this experiment gave further test to the non-referential hypothesis.

Besides these predictions, inferential analysis revealed no interaction effects in Experiment 5. With regards to the predictions drawn from the interference accounts, the data seems to tentatively support an attentional account, since we found a facilitation pattern in the regression path duration at the ADV region. This pattern however must be taken cautiously since it was not statistically significant. With regards to the non-referential hypothesis, the results from Experiment 5 suggest that spatial distance in the visual context does not modulate social relation interpretation if there is no connection between such visual context and the ensuing sentence.

4.4. Summary of results

In this chapter we examined two further questions about the relationship between spatial distance and abstract sentence interpretation. Specifically, in two eye-tracking reading studies, we first asked whether effects of spatial distance on reading times observed in three studies presented in Chapter 3 could extend to other semantic domains besides semantic similarity, namely to social relations. Secondly, we investigated whether these effects could also be found when no *and*-coordination of nouns matched the number of cards in the visual context. In Experiment 4, we found a reliable interaction effect between spatial distance and social relations

at the adverb region, however, with the opposite pattern relative to the first three experiments: longer first-pass reading times at the adverb region were found for sentences expressing friendly relations when cards were presented close together (vs. far apart), and the opposite for sentence expressing unfriendly relations.

Two accounts of interference effects, that could accommodate these results, were briefly discussed. Experiment 5 aimed to distinguish between these two accounts since they would predict different outcomes when words on cards are removed, but the sentences remain the same. Thus, experiment 5 was identical to experiment 4 with the exception that in all trials, cards in the visual context were empty. Inferential analyses showed no reliable interaction between spatial distance and semantic interpretation of social relations. Together, these results showed that spatial distance can modulate other abstract semantic domains (besides semantic similarity) such as social relations. However, the interaction between spatial distance and social relations evidence a different qualitative pattern compared to semantic similarity. Finally, and in contrast to semantic similarity, real-time interpretation of sentences expressing social relations seems not to be modulated by spatial distance depicted in a fully non-referential visual context.

CHAPTER 5

Visually-situated abstract language comprehension

5.1. Introduction

The study of visually situated language comprehension in real-time gained a lot of attention in the last two decades. A large number of studies have used the visual-world paradigm to study questions regarding syntactic, semantic and phonological processing in spoken sentence comprehension (see Henderson & Ferreira, 2004 for a review). Our review of existing evidence (Chapter 2) revealed that in spite of the vast amount of literature on language-visual context interactions, there are virtually no studies investigating the potential role of non-linguistic information for online abstract language comprehension. In contrast, evidence from experimental approaches, other than psycholinguistic, suggests that abstract conceptualization and visual information can in fact interact, affecting different behavioral measures.

In the literature reviewed, we identified key findings from both experimental psycholinguistics and cognitive psychology, which helped to delineate a number of research questions that remained poorly investigated. More specifically, the two first questions we asked were (1) whether abstract language interpretation could interact with visual information in real-time, and if so, what is the time course of these effects, and (2) whether language-

visual context interactions could be observed in the absence of a referential or lexical-semantic link, in what we called a non-referential visual context. These two questions were addressed in the first three experiments where we evaluated the effect of spatial distance on the interpretation of semantic similarity (Chapter 3). In two subsequent experiments (Chapter 4) we investigated two further questions with regards to the relationship between spatial distance and abstract language; first we asked (1) whether effects of spatial distance on abstract semantic interpretation could extend to other semantic domain (other than semantic similarity), namely social relations. Secondly, we asked (2) whether abstract semantic interpretation could co-index with spatial distance when the number of nouns in the sentence subject does not match with the number of objects in the visual context. Additionally, the second of these two experiments further tested the non-referential visual context hypothesis examined in Experiment 3.

This chapter is a discussion of the findings of this research project. The first section of the chapter is a summary of the key findings and the gaps identified in the literature review chapter. It also includes a discussion on the results of the five experiments, focusing on how the data contribute to answer the questions outlined in Chapter 1. In the second section we will contrast the present results with the prediction made by existing account of language processing in context, discussing the potential implication and challenges and suggesting possible updates for those accounts and their mechanisms, in order to be able to accommodate our findings. In the final section of the chapter, we will provide a summary of the main and more specific conclusions that could be drawn from the present research project. Additionally, we will comment on future lines of research we think could be valuable contributions to the understanding of (abstract and concrete) visually situated language comprehension.

5.2. General Discussion

5.2.1. Literature review chapter

Evidence discussed in Chapter 2, covered a wide range of language-world relations, and discussed different visual context effects on language comprehension. For example, we discussed how direct referential links to visually depicted objects (e.g., Tanenhaus et al., 1995; Spivey et al., 2002) and actions (e.g., Knoeferle et al., 2005, Knoeferle & Crocker, 2006) can modulate syntactic structuring and disambiguate references incrementally. We also discussed how semantic knowledge allows comprehenders to predict subsequent words of spoken sentences rapidly (Altmann & Kamide, 1999; Kamide et al., 2003). We saw that these effects could be observed even when visual information was no longer present on sentence onset (Altmann, 2004; Knoeferle & Crocker, 2007), and discussed the implications on the role of working memory and mental representation for the interplay between language and visual context. Finally, existing evidence suggests that language and visual attention interact when there is an explicit (referential) link, but also when subtler links connect visual objects with language, for instance lexical association and shared representational features (e.g., Huettig & Altmann, 2005, Huettig et al., 2006; Huettig & McQueen, 2007). As discussed in Chapter 2, these effects could originate from spreading activation of semantic, visual or phonological features, but could also reflect attempts to establish a reference.

Together the results showed that language and visual information can interact beyond direct referential links, which opens the possibility for abstract language processing to be studied. Nevertheless, the review revealed a lack of studies addressing potential connections between visual representations and abstract language comprehension (but see Duñabeitia et al. 2009). An existing theory (Paivio, 1986) assumes a clear distinction between the representation of concrete and abstract concepts in language. In this regard it predicts no relevant connections between abstract language and direct experiential modalities, such as

visual perception. Thus the study of the effects of visual information on abstract language comprehension could potentially be uninteresting. However, we discussed the possibility that the absence of such studies might be related to an experimental paradigm bias towards concrete language, based on the use of a (referential) linking hypothesis between gaze patterns and linguistic processing.

In addition, we discussed other existing theories, such as conceptual metaphor theory, (Barsalou, 2008; Lakoff & Johnson, 1999) that assume a more relevant role of direct experiential information for language comprehension, and reviewed empirical evidence showing that low-level visual information (e.g., orientation, movement, spatial information) can modulate sentence reading times as a function of the orientation or movement implied in the sentence (Zwaan & Taylor, 2006; Wassenburg & Zwaan, 2010; see also Kaschak et al., 2005 for end-of-sentence reaction times). These findings showed that task-irrelevant, unmentioned, visual information can rapidly and indeed incrementally modulate language interpretation (see also Coppens, Gootjes, & Zwaan, 2012), in coherence with theories of grounded cognition. And although theories of grounded cognition propose that experience is crucial for all levels of cognitive abstraction (see Barsalou, 2008; Lakoff & Johnson, 1999), we failed to find evidence in the existing literature showing that real-time interpretation of abstract sentences can be modulated by visual information.

We examined empirical evidence (beyond the language processing literature) addressing the conceptual metaphor hypothesis concerning the link between abstract concepts and perceptual representations. A number of studies found that visuospatial cues (e.g., brightness, location) can modulate reaction times and accuracy rates for categorization tasks (e.g., Meier & Robinson, 2004; Meier et al., 2004). Similarly, spatial schemas affected offline participants' responses to conceptual tasks (i.e., written answers) about time (Boroditsky, 2000). We then focused on two conceptual metaphors, namely SIMILARITY IS CLOSENESS and

INTIMACY IS CLOSENESS, revising experimental evidence showing how spatial distance and these two abstract concepts interacted in different behavioral tasks. These findings suggest that the link between abstract concepts and (visuospatial) experience proposed by conceptual metaphor theory goes beyond common linguistic expressions, and can modulate participants' responses in offline measures and response latencies.

Addressing the potential relationship between visual context and abstract language comprehension is a challenge not only for theories of grounded cognition but also for existing processing accounts of language in context. Both, processing models and grounded cognition theories, predict that visual information effects (or its representations in memory) on language comprehension should appear to be time-locked to the moment in the sentence in which the linguistic input drives the comprehender's attention to relevant aspects of the visual context (see Altmann & Kamide, 2009; Barsalou, Santos, Simmons & Wilson, 2008; Glenberg & Robertson, 1999; Kaschak & Glenberg, 2000; Knoeferle & Crocker, 2006, 2007; Taylor & Zwaan, 2008). Evidence suggests that co-indexing involved not just mapping between a word and a mental "picture"; the mental representation with which language co-indexes is a rich, multimodal and dynamic representation (see Altmann & Mirković, 2009).

In referential visual contexts, co-indexing between depicted objects or actions and language could occur through a referential mechanism (e.g., Tanenhaus et al., 1995). When the visual representations of a sentence are no longer present, linguistic information could co-index with visual representations present in memory (e.g., Altmann, 2004; Knoeferle & Crocker, 2007). When objects in the visual context are not mentioned, and thus a referential link is not obvious, it is less clear how this co-indexing occurs. As discussed above, lexical-semantic association, through a spreading activation mechanism, could potentially explain preferential looks to unmentioned target objects (see Huettig et al., 2006). We suggest that attempts to establish reference with a "good-enough" candidate in visual information could

also be relevant in the context of a visual-world experiment.

Virtually no studies had examined whether visual information can affect real-time interpretation of abstract sentences. Thus, there is no evidence showing how, and to what abstract language could co-index in the visual context. However, from our discussion it follows that a) co-indexing does not need a direct one-to-one lexical relation (the word-to-object relationship), but it can occur if at least some features are shared (e.g., Huettig & Altmann, 2005), and b) language can co-index to rather low-level visual information, such as movement or orientation (Zwaan & Taylor, 2006; Wassenburg & Zwaan, 2010). With these two ideas in mind, and drawing predictions from conceptual metaphor theory, we hypothesize that if visual information can modulate abstract language interpretation incrementally, it will do so through co-indexing low-level aspects of visual context and a critical linguistic input in the sentence, which would drive participants' attention to the visual representation.

5.2.2. Empirical chapters

We conceptualized the gaps we found in the literature in four research questions. To recap, our research questions were;

- I.** Can abstract language interpretation interact with visual information in real-time, and if so, what is the time course of these effects?
- II.** Can language-visual context interactions be observed in the absence of a referential or a lexical-semantic link?
- III.** Does the effect of spatial distance on real-time semantic interpretation extend to different semantic domains, such as similarity and social relations?
- IV.** Does co-indexing between abstract semantic interpretation and spatial distance depend on whether the linguistic element (i.e., the subject noun of the sentence) can be mapped in elements in the visual context?

The first three eye-tracking reading experiments assessed the first two research questions. We manipulated spatial distance in the visual context (playing cards close together vs. far apart), and the semantic content of written sentences (similarity vs. dissimilarity between two abstract nouns). Spatial distance and semantic similarity were never referentially or lexico-semantically linked¹⁰, and between experiments, we varied their relatedness. In Experiment 1, cards on critical trials showed two abstract nouns that re-appeared in the ensuing sentence. In Experiment 2, cards were blank before each critical sentence but participants learnt the pairing of abstract nouns in blocks, prior to reading the sentences, and cards on most filler trials showed words that re-appeared in the following sentence. In Experiment 3, only blank cards were presented, thus eliminating any overt relationship between card distance and the semantic similarity of the sentential nouns.

In Experiment 1, reading times were faster when sentences about similarity were preceded by cards-with-words close to each other (vs. far apart), and when sentences that expressed dissimilarity were preceded by cards-with-words far apart (vs. close together). This gaze pattern was reliable in first-pass reading times at the ADJ, the VP2, and NP3 regions. In Experiment 2, we observed the same reading time pattern but in later measures (i.e., total reading time). Spatial distance effects on the interpretation of semantic similarity thus appeared even when critical trials did not present words in the visual context but empty cards instead. Together the first two experiments showed that when visual context was related to an ensuing sentence (through words on cards for critical items in Experiment 1; through words on cards for filler items in Experiment 2), then spatial distance facilitated semantic interpretation of similarity between two abstract nouns.

Results from Experiment 3 replicated rapid incremental spatial distance effects even in the absence of any overt relationship between the cards and the subsequent sentence. When

¹⁰ It could be argued that the repetition of words (on the cards and in the sentence) counts as a “referential” link. Note, however, that merely relating individual words on the cards to words in the sentence could not have caused the observed spatial distance effects. Instead, for these effects to emerge, participants must process the distance between two objects, the semantic similarity between two abstract nouns, and then integrate these two.

the visual context showed two blank cards at different distances, spatial information rapidly influenced incremental semantic interpretation of sentences that expressed similarity or dissimilarity between two abstract nouns. First-pass times at the NP2 region for sentences that expressed similarity were shorter when preceded by cards close together compared to far apart and first-pass times for sentences that expressed dissimilarity were shorter when cards were far apart compared to close together.

The results from these three experiments revealed that non-linguistic visual information can rapidly and incrementally modulate the semantic interpretation of abstract sentences; and that this can happen even in the absence of referential or lexical-semantic links. However, there are some differences in the time course of these effects between the experiments. While in the first two experiments we observed the first reliable interaction effects emerging at the adjective region (where similarity was made explicit), spatial distance effects appeared earlier in Experiment 3, namely at the second abstract noun (where similarity was implied). Conversely, in Experiment 3 no reliable interaction effects were observed at the adjective but only in the subsequent region (the verbal phrase). Furthermore, Experiments 1 and 3 revealed effects in the earliest measure reported (first-pass reading times) whereas in Experiment 2, spatial distance effects were delayed, appearing in total times. We will briefly discuss these open questions and provide with tentative explanations for them.

In the first two experiments, participants always read the two first sentential nouns before reading the critical sentences, either presented on the cards preceding the sentence (Experiment 1) or in a list of word pairs before a block of trials (Experiment 2). We argue that this previous processing of the noun pairs could have discourage deep semantic processing of the second noun embedded in the sentence. Elsewhere in the literature, it has been shown that word repetition facilitate word recognition and categorization (see, e.g., Jacoby, 1983). If deep semantic processing is necessary for spatial distance effects to modulate semantic processing in real time, it is plausible that when participants process the second noun

superficially, such effects are not observed. In Experiment 3, by contrast, participants encounter the second noun for the first time only after seeing cards moving in different directions. This would possibly demand a deeper semantic processing of that sentence region and integration with preceding (linguistic and non-linguistic) context, resulting in the emergence of early effects.

The reasons why spatial distance effects were not observed at the adjective region but only after that in Experiment 3 are undoubtedly more intriguing. We can only speculate that the integration of semantic similarity implied at the second noun, and non-referential spatial distance in the visual context could prompt participants to move their internal attention towards the visual representation as they continue with the sentence. This might have produced a different effect of spatial distance at the adjective region. Interestingly, the reading times pattern for this region (in regression-path duration), showed essentially the opposite pattern compared to that at the second noun; reading times for similarity-conveying sentences were longer when cards were close together than far apart. This pattern of effects, however, was not statistically reliable, weakening this argument.

Finally, delayed effects observed in Experiment 2 could be related to a difference in the experimental procedure, which could have turned to be more demanding in terms of processing. In Experiment 2, participants were asked to read and learn a number of word pairs before performing a block of experimental trials. As in Experiment 1, participant had access to the pairs of words later embedded in the sentences; however, they had to retrieve this information from working memory instead of having it together with the visual context. This potential integration of visual context, learnt pairs and sentence meaning might have resulted in higher processing difficulties, delaying the influence of spatial distance to later measures. An alternative, all these differences could have emerged because the precise time course of non-referential visual context is somehow variable and the effects are not strongly robust. Future research might clarify these differences.

Despite these differences in the time course of the effects, together the results of these three experiments revealed that non-linguistic visual information can rapidly and incrementally modulate the semantic interpretation of abstract sentences and that this can happen even in the absence of referential or lexical-semantic links. Together the results from these three experiments are the first evidence for rapid effects of non-referential spatial information on the incremental comprehension of abstract language. They broaden previous findings of spatial distance effects on similarity ratings and response times (Boot & Pecher, 2010; Casasanto, 2008) to real-time language processing. In agreement with theories of grounded cognition, they provide strong evidence for the role of spatial information in abstract language comprehension. These findings are further compatible with the view that the metaphorical mapping mechanism proposed in conceptual metaphor theory has implications for incremental language processing.

Conversely, there is some evidence for rapid effects of abstract language on visual attention during spoken language comprehension. When participants heard an abstract word (e.g., “smell”), they inspected objects (e.g., nose) associated with that word more often than non-associated objects (Duñabeitia et al., 2009). Other studies have demonstrated that visual context and concrete language can interact rapidly when objects in the visual context are not mentioned, for instance, through subtle associations such as the shape of an object (e.g., Dahan & Tanenhaus, 2005; Huettig & McQueen, 2007). Those findings, unlike the present ones, can be accommodated by a lexical-associative account (e.g., relating the shape of a mentioned object to that of similarly shaped unmentioned objects). The new insight from our results is that participants processed both the spatial distance between playing cards and the semantic similarity of two unrelated abstract nouns, and integrated the resulting representations during incremental semantic interpretation.

Spatial distance effects were observed at the adjective region of the sentence where similarity (or dissimilarity) was explicitly conveyed, and extended to subsequent regions of

the sentence. Furthermore, these effects were also found at the second noun phrase, which could imply semantic similarity only in the context of being preceded by a synonym (when implying similarity) or an antonym (when implying dissimilarity). Consequently these results cannot be accommodated by a lexical-semantic account, since they require compositional processing. One aspect, however, that this set of experiments could not tease apart is whether co-indexing between spatial distance and semantic similarity happens directly or through the co-indexing between the two nouns of the *and*-coordination and the two objects in the visual context. Moreover, it was unclear whether the effects extend to other abstract domains, such as social relations. In two further experiments we examined these two open issues.

As discussed in Chapter 2, both behavioral (Matthews & Matlock, 2011; Williams & Bargh, 2008a) and non-experimental (Lakoff & Johnson, 1999) evidence suggested that spatial distance could be systematically related to the way people understand social relation in language. Thus, we investigated whether spatial distance effects could be observed on real-time semantic interpretation even when no coordination of nouns matched the number of objects in the visual context, but instead co-indexing spatial distance directly with the abstract notion of friendliness. Using the same paradigm described in Chapter 3, we manipulated spatial distance between two playing cards and in successive experiments we changed the relationship between those cards and an ensuing sentence. In Experiment 4, the subject-noun and the object-noun of each sentence were presented each on a card (see Experiment 1), while in Experiment 5, cards showed no words at all (see Experiment 3). Critical sentences expressed a friendly (vs. an unfriendly) interaction between the two characters, which corresponded to the sentence subject and object.

When the cards presented two first sentential nouns (Experiment 4), a reliable interaction was observed between spatial distance and social relation at the predicted region, and crucially the effect appeared in first-pass reading times. Finding such effects only after the critical adverb region (or even at the adverb region but only in a later measure, e.g., total

reading time), could have been interpreted as evidence for co-indexing between nouns in the sentence and objects in the visual context, rather than spatial distance being co-indexed with the adverb. These findings support a view in which abstract concepts can be rapidly co-indexed with spatial information. However the interaction pattern was the opposite of that observed previously.

To accommodate this result, two accounts of interference vs. facilitation effects, and their predictions for a subsequent experiment, were presented. While both accounts could accommodate the interference-like pattern observed in Experiment 4, an account focused on the integrability of semantic content and perceptual information would predict an interference-like effect, while an account focusing on attentional modulation would predict facilitation-like patterns. To test these predictions, in Experiment 5, we removed all words from the cards in the visual context, however, no reliable interaction between spatial distance and social relations was observed. These findings suggest that a fully non-referential visual context does not always modulate online semantic interpretation of related abstract concepts. But why does spatial distance in a non-referential visual context modulate semantic similarity interpretation (see Experiment 3), but does not modulate interpretation of social relations? When comparing the two experiments presenting related visual context (two sentential nouns in the visual context relating directly to the ensuing sentence) a clear difference can be seen in the extent of the effects. While in Experiment 1 (semantic similarity) the effects of spatial distance began at a critical region and extended to the end of the sentence (in the three subsequent regions), in Experiment 4 (social relationships) the effect appeared at the critical region only to vanish afterwards.

In those experiments that examined semantic similarity sentences, we observed a decrease in the extent of the effect when cards on the visual context contained no words; interaction effects were observed at the critical region, to re-appear only once after the

critical adjective. Taking this into account, it is possible that although the critical manipulation in the visual context is never directly referred to in the sentence, the fact that the visual context somehow relates to the sentences strength the non-referential spatial distance effect on semantic interpretation. Potentially, if the effect is short-lived even when there is a relation between visual context and the sentence, the predicted effect is more unlikely to be observed with a completely non-referential visual context.

An alternative reason why the effect of spatial distance could be weaker for social relation compared to semantic similarity is that in contrast to semantic similarity, social relations have a larger network of associated perceptual representation that can be activated during online sentence comprehension. Indeed, social relations can be expressed in terms of temperature (“they gave me a warm welcome”) or verticality (“the highest position one can aspire in the company”) in addition to spatial distance, while semantic similarity seems to rely mostly in spatial distance (see Lakoff et al., 1991). It is possible that the interpretation of semantic similarity systematically (if not exclusively) activates spatial distance representations, while social relation do so less systematically, sometimes activating other perceptual representations (e.g., temperature, verticality) as the comprehender constructs the meaning of the sentence. The strength of association between the semantic domain and spatial distance may predict the extent of the effects of perceptual representation and interpretation of a given semantic domain. Another possibility to be considered is the fact that there is no noun coordination to match the two objects in the visual context. Perhaps the co-indexing is stronger when the number of objects matches the number of nouns in the subject of the sentence. This is also plausible based on the comparison between Experiments 1 and 4, where a less extended effect was evident in the latter. These three possibilities are not mutually exclusive, so it is still even possible that each of the explanations contribute to the null effect found in Experiment 5.

5.2.3. Interim summary

The experiments presented in this thesis provided a set of novel findings about the relationship between visual information and abstract language comprehension in real-time. The first three experiments provided evidence to answer questions I and II. First, abstract language was modulated by non-linguistic information in the visual context. Furthermore, even in the absence of clear referential or lexical-semantic links, spatial distance put across in the visual context distinctively modulated real-time semantic similarity interpretation. Results from Experiment 4 allow us to give a tentative answer for questions III and IV. Spatial distance effects extended to the interpretation of social relation. Furthermore, these results showed that a match between objects in the visual context and linguistic elements (e.g., nouns) in the sentence, is not a condition of co-indexing between abstract language and low-level visual information such as spatial distance. To the best of our knowledge, these are the first results showing that non-referential spatial distance can modulate incremental interpretation of semantic similarity and social relations.

Results from Experiment 5, however, cast some doubts on the extent to which non-referential visual context can influence sentence comprehension rapidly and incrementally. The null effect observed in the last experiment suggests that in contrast to semantic similarity, other abstract semantic domains such as social relation are not modulated by fully non-referential visual context. We outlined at least three potential explanations for the absence of fully non-referential spatial distance effects on real-time interpretation of social relation. Further research in this direction, could clarify whether the effects of fully non-referential visual contexts can be observed for social relation interpretation, for instance, when a coordination of nouns matches the number of objects in the visual context.

5.3. Implication for models of situated language processing

In the previous section we summarized the evidence that allowed us to create a framework that could inform our hypothesis with regards to the relationship between non-referential visual information and abstract language. We then enumerated the research questions that emerged from that review and described the results of five experiments, focusing on the extent to which our findings answer those research questions. However until now we have not explained how these results might be relevant for accounts of situated language comprehension. In Chapter 2, we discussed two models of situated language comprehension; the CIA (Knoeferle & Crocker, 2006, 2007) and an affordance-based account (Altmann & Kamide, 2007). We discussed their central mechanisms and how evidence that non-referential visual context affects abstract language comprehension in real-time would require their revision. We think that the results presented in this thesis are not incompatible with such accounts, but they present a number of challenges in terms of the mechanisms by which visual information interact with language comprehension and with regards to the range of real world-language relations that these models can account for. In this section we will address that issue.

Existing accounts of visually situated language comprehension (e.g., Altmann & Kamide, 2007; Altmann & Mirkovic, 2009; Knoeferle & Crocker, 2006, 2007; Tanenhaus et al. 2000) have described models in which language interpretation and visual attention are closely coupled in time, such that unfolding stream of linguistic information can rapidly direct attention to objects and actions presented in a visual scene or their derived mental representation in memory. Our results are in line with these predictions. Spatial distance between cards modulated reading times in regions in which semantic similarity or friendliness was expressed or implied in the sentence. In contrast, some of these accounts rely on a referential linking hypothesis between language and objects or concrete events depicted

visually (Knoeferle & Crocker, 2006, 2007; see also Tanenhaus et al. 2000), and they are not equipped with mechanisms that can accommodate effects of non-referential visual contexts.

An affordance-based account (see Altmann & Kamide, 2007) seems to be able to accommodate preferential and even predictive looks to objects that not yet mentioned (or not mentioned at all). However, that account still relies on mapping between language and the event that objects in the visual context can afford. Therefore, although it explains how looks to unmentioned objects take place, it still needs a context in which a referential visual scene enables a comprehender to make inferences and predictions with regards to the incoming linguistic input. On the contrary, our results rely on a much subtler relationship between language and visual information, which is fundamentally different from a referential link. Distance between objects in the visual context was never directly referred to. Indeed, those experiments in which cards presented written words that reappeared in the subsequent critical sentence (experiment 1 and 4) might be have enhanced the mapping between language and the visual context. Yet, such a link was not a condition for the effects of spatial distance on abstract language interpretation (Experiments 2 and 3, but see experiment 5).

Moreover, current accounts of visually situated language processing are still unprepared to explain visually situated abstract language processing. As discussed in Chapter 2, psycholinguistic research has not systematically explored the potential relationship between visual information and abstract semantics. While none of these accounts have explicitly rejected the possibility that visual information could rapidly and incrementally modulate abstract language, neither have they explicitly included such a possibility. Their focus on a referential link or on affordances has, from our point of view, obscured a potential wider spectrum of language-world relations, including how abstract language might be affected by low-level visual information. Referential accounts require a clear depiction of objects and events described by language, and affordance accounts need objects and situations with a restricted set of affordances that subsequently can be related to language and inform

comprehension. Consequently, language that is not about concrete objects or actions cannot be easily related to visual contexts under these views. In this sense, our results present a challenge to these accounts.

We argue, nevertheless, that the mechanisms by which visual information affects abstract language comprehension are not fundamentally different from those proposed by Knoeferle and Crocker (2007), and Altmann & Kamide (2007). These accounts provide online mechanisms that explain how language guides (visual or internal) attention in the cognitive system towards the aspect of the visual context that are informative for the linguistic interpretation and prediction generation. Yet, in order to accommodate a wider range of visual world-language relations, some extensions and refinements on the original versions of these accounts are needed. In addition, we will introduce a novel mechanism borrowed from conceptual metaphor theory, which enables us to make clear prediction with regards to what kind of visual information could be relevant for abstract language. The metaphorical mapping mechanism also extended the kind of possible mapping between visual information and language processing, to relationships that have not been yet explored by any processing accounts of situated language.

5.3.1. Visually situated abstract language comprehension

In this section, we will discuss the cognitive mechanisms by which low-level visual information (i.e., spatial distance between objects) rapidly and incrementally modulates the comprehension of abstract sentences about semantic similarity and social relations.

5.3.1.1. Metaphorical mapping

The first mechanism relevant to this relationship is metaphorical mapping (Gallese & Lakoff, 2005; Grady, 2005; Lakoff & Johnson, 1980, 1999). According to Lakoff & Johnson

(1980), metaphorical mapping is fundamental to the construction of meaning and conceptual understanding. This mechanism is particularly important for abstract concepts (see Grady, 1997), such as similarity or intimacy, because, in principle, they lack unequivocal physical correlates. What metaphorical mapping does is to link “our sensory-motor experience to the domain of our subjective judgments” (Lakoff & Johnson, 1980: p. 256). For example, situations in which people experience psychological affection are strongly correlated from early childhood with the physical experience of warmth, as they are held close to their caregivers. Likewise, emotional intimacy is often experienced in the context of physical closeness with others. According to conceptual metaphor theory, such experiences allow people to understand affection and intimacy at a more abstract representational level and give origin to conceptual metaphors such as AFFECTION IS WARMTH and INTIMACY IS CLOSENESS (Lakoff & Johnson, 1999; see Williams & Bargh, 2008a, 2008b for experimental evidence). These types of recurrent experiential correlations are the basis of conceptual metaphors (Gallese & Lakoff, 2005; Grady, 2005). In this sense, metaphorical mapping might be understood as a mechanism relevant for conceptual learning (see Grady & Johnson, 1997; Johnson, 1999). But is it metaphorical mapping a mechanism of incremental language comprehension?

As has been discussed in previous chapters, converging evidence suggests that links between perceptual representation, such as visual information (or temperature for that matter; see Williams & Bargh, 2008b), and abstract conceptualization can be observed experimentally in a number of behavioral studies. Yet, there is little evidence for the relevance that these links could have during online sentence comprehension. We made use of those links proposed by conceptual metaphor theory to investigate whether abstract language comprehension could be affected by a non-referential visual context. The results from our experiments are the first evidence suggesting that metaphorical mapping—as a grounding mechanism by which abstract concepts are related to direct experience—has implications for real-time language comprehension. This finding, however, seems to be more robust for the

conceptual metaphor SIMILARITY IS CLOSENESS, compared to INTIMACY IS CLOSENESS. Thus, based solely on our data we cannot unequivocally state that metaphorical mapping is an active mechanism during real-time language processing. We argue, consequently, that metaphorical mapping could be responsible for the initial link between perceptual experience and abstract concepts, but it is still unclear whether this sort of mapping is re-created every time we understand abstract sentence in real-time.

5.3.1.2. Co-indexing

In contrast, there is another mechanism that has previously been assumed to be involved in connecting visual representations and language comprehension, in real-time. We argue that a co-indexing mechanism (see Knoeferle & Crocker, 2007) might play an important role for the integration of abstract language and non-linguistic visual information. Initially, co-indexing was assumed to integrate a narrow set of visual representations (namely, objects and actions) to linguistic interpretation by linking linguistic meaning, visual cues (e.g., object location), and world-knowledge (Knoeferle & Crocker, 2007, p. 541). In the context of a referential visual context, where concrete objects and actions are visually depicted and verbally described, partial language interpretation guides visual attention to objects or actions in the scene. As reference is established, the linguistic input and its referent in the visual context are co-indexed, allowing visual information and world-knowledge to be integrated to language comprehension incrementally. Thus, co-indexing (in contrast to metaphorical mapping) is clearly time-locked to segments of the sentence in which linguistic meaning directs attention to relevant aspects of the visual context in real-time.

Importantly, Knoeferle and Crocker (2007) showed that co-indexing might occur between language and a co-presented visual context, as well as its mental representation in working memory. As discussed in Chapter 2, spoken language can guide visual attention to the location where objects that are referred to previously were (e.g., Altmann, 2004). The co-

indexing mechanism is highly specified with regards to its content. As presented in the CIA, it links nouns with objects and verbs with actions (Knoeferle & Crocker, 2007). However, other findings in the literature could be taken as evidence for subtler (perhaps non-referential) co-indexing between language and visual information. For example, the orientation of an object in a picture (e.g., a vertical as opposed to a horizontal toothbrush) and the sentential prepositional phrase that implies orientation of an that object (e.g., "...the toothbrush in the cup vs. in the sink..."; see Wassenburg & Zwaan, 2010; Coppens et al., 2012), could potentially be co-indexed in such a way that toothbrush is co-indexed with the object itself.

When language refers to objects in a (co-presented or previously seen) visual context, co-indexing builds a connection between the visual referent, the referring linguistic expression, and the world-knowledge associated with the referent. Does this imply that co-indexing requires reference to be established before it can create the link between all these sources of information? According to Knoeferle & Crocker (2007), reference resolution and co-indexing, as well as other key processes relevant for language-visual context interaction "in reality partially overlap and occur in parallel" (p. 540). This implies that co-indexing could potentially work independently of whether the objects or events depicted in the visual context are referred to. As in the above-mentioned example, the sentence "...the toothbrush in the cup" does refer to a toothbrush. However, it does not refer (in a strict sense) to the orientation of the toothbrush. Instead, the prepositional phrase implies the orientation of the toothbrush. Crucially, evidence from eye-tracking reading times (Wassenburg & Zwaan, 2010) and event-related potentials (Coppens et al., 2012) show that visual information effects appeared precisely at the prepositional phrase. This interpretation of those results is in agreement with the linguistic focus hypothesis (see Taylor & Zwaan, 2008, Zwaan, Taylor & de Boer, 2010), and it further provides an explicit mechanism for the influence of subtle aspects of the visual context on concrete language comprehension and when there is a least one object referred to.

But how does the co-indexing mechanism create a link between a fully non-referential visual context and abstract language comprehension? In Chapter 4, we discussed two options in which co-indexing could happen when two objects are presented in the visual context and the subsequent sentence has an *and*-coordination of nouns as sentence subject (Experiments 1, 2 and 3). As a first possibility, each noun phrase could be directly co-indexed with each of object of in the visual context (i.e., the playing cards). This is especially likely when the cards presented the two first sentential nouns in the visual context (Experiment 1). It could be also the case, when no words were presented in the visual context, since the number of nouns in the subject of the sentence matched the number of objects in the visual context. As participants encounter the coordination of nouns in a sentence such as ‘*Peace and war are certainly different...*’, they might link the meaning of each noun (“peace”, “war”) to each card. Additionally, word meaning (embedded on its linguistic context) activates long-term semantic associations (or world-knowledge), such as “war” being the opposite of “peace”. World-knowledge and the spatial representation derived from the array of cards (i.e., spatial distance between them) would be co-indexed, even if the objects in the visual context are not referred to.

A second possibility is that the spatial representations of distance between objects might co-index with the abstract semantic representations of similarity, implied by the noun and-coordination and the adjective. In this case, participants encounter the coordination of nouns in a sentence such as ‘*Peace and war are certainly different...*’, but the nouns are not co-indexed to playing cards. Instead, processing the second abstract noun activates world-knowledge implying semantic similarity (or dissimilarity), which co-index with relevant aspects of the visual context (spatial distance), and word meaning (more similar to the prepositional phrase in “the toothbrush in the cup” example). None of these two options require a referential link. On the other hand, it is arguable that the first option is more likely to happen when cards-with-words (Experiment 1) are presented in the visual context, while the

second explanation might be more plausible when blank cards are presented in the visual context (Experiment 3). Clearly, these two options are not necessarily mutually exclusive, yet it was important to explore the possibility that co-indexing might happen beyond reference, since this is not predicted by the CIA. Whichever is the case, the first set of studies could not distinguish between the two options. Therefore, whether co-indexing requires that objects in the visual context (e.g., playing cards) are linked to content words (e.g., nouns) in the sentence was still unclear.

This open issue was addressed by modifying the structure of critical sentences (Experiments 4 and 5), such that the first noun appears as the subject of the sentence, and the second noun as the sentence object (e.g., '*Sandra met cheerfully her aunt at the health center*'). In addition, we extend the investigation of non-referential visual context to other semantic domains. In these experiments, the critical region appeared in between the two sentential nouns. The visual context was the same as in Experiment 1 (for Experiment 4) and Experiment 3 (for Experiment 5). If co-indexing requires a match between cards and nouns before spatial distance can affect semantic interpretation, spatial distance effects should appear only after the second noun phrase (*'her aunt'*) has been encountered, even though the critical adverb region (*'cheerfully'*) precedes it in the sentence. Alternatively, if spatial distance can co-index directly with abstract adjectives/adverbs, then effects should appear before the second noun phrase is encountered, namely in first-pass reading time or regression path duration at the critical adverb region.

The results from the last two experiments are somewhat inconclusive. We found an interaction effect in Experiment 4 (when cards and sentence were related via words on cards), but we did not find any effects in Experiment 5. Nevertheless, the reliable interaction effects found at the adverb region in first-pass reading time (Experiment 4), tentatively suggest that spatial distance can co-index directly with abstract adverbs before objects in the visual context and nouns in the sentence could be matched. The results from the last experiment,

however, weakened this finding, and suggest that, although observable, that effect is not that robust. Together, the results from five experiments suggest that a co-indexing mechanism could be playing an important role in mediating the influence of the effect of spatial distance on abstract semantic interpretation.

The two mechanisms described so far could potentially explain two important aspects of the presented findings, namely the origin of the connection between spatial distance and semantic similarity/social relations, and how the two sources of information are integrated to inform real time language comprehension. Metaphorical mapping and co-indexing allow us to make clear predictions about the time course of perceptual representation effects on abstract sentence comprehension. However, are these two mechanisms sufficient to explain the pattern of the results of the present experiments? To answer this question it is important to contrast some key differences between the experiments in which co-indexing was originally proposed and the present study. In Knoeferle & Crocker (2007), the visual context acts as a constraint for the syntactic structuring of sentences. Thus, while the target sentences are locally ambiguous, there is only one interpretation that matches the previously observed depicted event. In contrast, in our studies, the visual context does not constraint sentence interpretation for two main reasons; on one hand, critical sentences are unambiguous, thus there is only one possible syntactic interpretation. On the other hand, a non-referential visual context does not visually depict the situation described in the subsequent written sentence; therefore it does not provide a true (or false) value that could constrain semantic interpretation.

We argue that co-indexing could link informational sources in both experimental conditions (when visual context is a constraint and when it is not). Yet, the consequences of co-indexing for language interpretation are potentially quite different under these two conditions. In the case that the visual context is a visual depiction of the subsequent sentence (constraining interpretation) visual information constantly triggers a revision and reconciliation (whenever necessary) of partial linguistic interpretation. By contrast, when the

visual context does not constrain sentence interpretation, visual information does not prompt a revision or reconciliation of current interpretation. We argue that instead, non-referential visual information effects on abstract language might be working at more basic level mechanism such as spreading activation between related representations. Priming is generally used to refer to a processing advantage of a target-stimulus (e.g., target-word) whenever a related stimulus (e.g., prime-word) has previously been processed (e.g., Anderson & Pirolli, 1984; Collins & Loftus, 1975). In other cases, priming refers to the repetition or processing advantage of a given syntactic structure when a similar sentence structure has recently been processed (see Pickering & Ferreira, 2008 for a review). Here we use the term priming in a wide sense, without committing ourselves to any specific processing level, such as word, clause, sentence or even a picture (cf. Coppens et al., 2012; Connell & Lynott, 2012). We recognize however that spreading activation is a low-level mechanism that could be involved in various levels of processing.

5.3.1.3. Spreading activation beyond semantic priming

5.3.1.3.1. Facilitation effects

The present set of studies could be described as presenting a prime (spatial distance), which we predict should affect the processing of a target (critical regions of a sentence conveying similarity or social relations). In this sense, our experimental paradigm works as a priming paradigm. We also acknowledge that the observed effects are likely to be of a semantic nature (since we manipulated the semantic content of critical sentences). Evidence from previous studies discussed in Chapter 2 have already shown that visual context and language (both concrete and abstract) can rapidly interact when objects in the visual context are not mentioned, but when they are associated with a target word (e.g., smell – nose, see Duñabeitia et al. 2009). Even subtler lexical associations, such as the stereotypical color or shape of objects, have shown to rapidly guide comprehenders attention towards unmentioned

objects in the visual context when they shared the color or shape of the referred concept (e.g., Dahan & Tanenhaus, 2005; Huettig & Altmann, 2011; Huettig & McQueen, 2007). These findings, both for concrete and abstract language, go beyond referential links and yet they can be accommodated through a lexical-associative account (or the spread of activation of single word semantics). Such effects are triggered by a specific lexical input and they seem to be insensitive to contextual appropriateness (see Huettig & Altmann, 2007), which would require compositional processing. These results, which can be characterized as lexical-semantic priming effects, are indeed different from ours.

But how are our results different from lexical-semantic priming? Critically, what makes a simple semantic priming account insufficient is the time course of the reported effects. The results of our studies showed that the visual representation of spatial distance between playing cards distinctively modulates abstract semantic interpretation during reading at critical regions conveying semantic similarity or intimacy. Even if participants verbalized the distance between cards (e.g., by silently saying “close” or “far”), the regions of the sentences in which we observed effects of spatial distance (e.g., the second noun “war”, in Experiment 3 and adverb “gently” in Experiment 4; see the Appendix for sentence examples) were not associated with such lexical entries. Furthermore, these effects extend to subsequent sentence regions not related to spatial distance. To accommodate these results spreading activation must have happened detached from specific lexical items. Those critical regions, embedded in their particular sentence context, implied specific semantic content, such as similarity in the case of the and-coordination of two abstract words (Experiment 3), or intimacy in the case of the abstract adverb (Experiment 4). The observed effects therefore required a compositional process.

The spreading-activation construct has been proposed to explain several aspects in a single-word prime context (e.g., Anderson & Pirolli, 1984; Collins & Loftus, 1975; Hutchinson, 2003), and also in the context of real-time sentence comprehension, when a

specific lexical item in the sentence (the prime) precedes another specific lexical item (the target; e.g., Carroll & Slowiaczek, 1986; Morris, 1994). The effects found in our studies certainly arise from processing different semantic content (e.g., similar vs. dissimilar) after processing different contextual information (i.e., cards far apart vs. close together). Likewise, it appears to be clear that the observed effects happened beyond the lexical level. Elsewhere, priming has also been proposed to work in a more abstract way and beyond the word-form level (see, e.g., Thothathiri & Snedeker, 2008; Pickering & Branigan, 1998) although in the context of syntactic priming during comprehension or production. We argue that, similarly, the effects of spatial distance on sentence processing reported here results from spreading activation of perceptual representation onto conceptual processing. In fact, the notion of spreading activation is critical for conceptual metaphor theory and metaphorical mapping (see Lakoff, 2008, 2012).

As discussed earlier, if metaphorical mapping is to consolidate—that is, if an abstract concept is to be grounded in concrete experience—both the concrete experience (e.g., spatial distance) and the abstract representation (e.g., similarity, intimacy) need to repeatedly occur together. According to Lakoff (2012), the association between perceptual and abstract representations is based on a Hebbian learning principle (Hebb, 1949). This principle has been also assumed to be responsible for the activation of specific sensorimotor cortical areas during comprehension of concrete words (see, e.g., Pulvermuller, 2005, 2013). When cortical areas, responsible for processing perceptual properties for objects or motor programs for actions, are activated in a synchronized way with cortical areas responsible for processing words referring to such objects or actions, these two potentially different neural groups start to connect to one another. If this synchronized activation occurs repeatedly and recurrently, their connection becomes stronger (see Pulvermuller & Fadiga, 2010). Similarly, when perceptual processing (e.g., perceived distance) is persistently recruited and activated during

abstract concepts acquisition, cortical areas responsible for processing both representations will be strongly interconnected.

As discussed in Chapter 2, this notion is not new for processing accounts of situated language comprehension. In fact, the affordance-based account strongly relies on feature overlap and spreading activation mechanism between similar representations (see Altmann & Kamide, 2007). If perceptual and abstract representations are strongly connected, this would have a direct consequence on processing. As Collins and Loftus wrote in their original presentation of a spreading-activation theory of semantic processing, “when a concept is processed (or stimulated), activation spreads out along the paths of the network” (1975; p. 411). This would mean that when spatial distance is processed, activation spreads to representations associated with it. Such a mechanism, in conjunction with co-indexing, can explain the facilitation or priming effect we observed in the first three experiments.

5.3.1.3.2. Interference effects

One important challenge for this thesis is to provide an account that could potentially accommodate both facilitation (Chapter 3) and interference effects (Chapter 4) of spatial distance on sentence reading times. The results we observed in Experiment 4, however, were weaker than those found in the first three experiments. They are also less extended, appearing only at the critical word. Moreover, they seem to be much less robust, since they were not replicated in a fully non-referential visual context. Thus, the results from Experiment 4 must be taken more cautiously. Nevertheless, and as discussed in Chapter 4, the interference effect between perceptual and related linguistic representations have previously been reported in the literature (see Connell, 2007; Kaschak et al., 2005; Richardson et al., 2003). Therefore, a complete account of the interaction between visual information and abstract language should at least contemplate this possibility.

Metaphorical mapping and co-indexing contribute to explain how a particular spatial representation relates to abstract language, and how during incremental language processing, the mental representation of this spatial representation is integrated with partial linguistic interpretation and world-knowledge. Moreover, a spreading activation mechanism can explain how attention moves from stored-knowledge activated by the linguistic input to the spatial representation in working memory, priming reading times for concepts that are associated. A critical question that remains open is whether this mechanism, together with metaphorical mapping and co-indexing, could also accommodate the reversed pattern we observed in Experiment 4. Existing evidence suggests it could.

The theoretical framework in which the affordance-based account is embedded, assumed that semantic representations (or concepts) are graded structures composed by parts or features, which in turn are shared with other concepts (e.g., Cree & McRae, 2003; McRae, Cree, Seidenberg & McNorgan, 2005; Rogers & McClelland, 2004, 2008). Moreover, in accordance with this view associated and similar representations are activated in parallel and compete for attention (e.g., McClelland & Rumelhart, 1981). At the word level, concepts that shared features are commonly denominated neighbors. For instance, the word “dog” is a semantic neighbor of the word “cat”, since they share semantic features such as is-animal, has-four legs, can-run, etc. Similarly, the words “candle” and “candy” are phonological neighbors, since they are phonologically identical until the middle of the second syllable. Critically, there is a vast amount of evidence that neighbor relationships can both elicit priming and inhibitory effects in a variety of linguistic tasks (see Chen & Mirman (2012) for a short but detailed review).

Recently, Chen and Mirman (2012) proposed that facilitatory and inhibitory effects of neighbors are governed by the same mechanism. Like others, (e.g., McClelland & Rumelhart, 1981) they assumed that parallel activation and competition require inhibitory mechanisms. However, “weakly active word units had very little inhibitory effect on other word units, and

strongly active words units had a very strong inhibitory effect on other word units” (p. 420). In a number of simulations, they provided evidence that “weakly active representations facilitate semantic processing while strongly activate perceptual representations interfere with it.” (p. 422). Consequently, they concluded that whether parallel activation results in facilitation or interference is determined by the difference in the degree of activation between neighbors. Therefore, there is evidence that concepts that are related through features can trigger both facilitation and interference effects. This opens the possibility that perceptual representation could also elicit both facilitation and interference on abstract sentence comprehension.

In addition, a recent account has been delineated by Connell and Lynott (2012), with the purpose of accommodating conflicting results from studies examining the role of perceptual representations during conceptual processing (e.g., Richardson et al., 2003; Kaschak et al., 2005; Bergen et al., 2007). This account was briefly introduced in Chapter 3. The authors presuppose that perceptual and the attentional systems are strongly entangled, and thus modality-specific attentional resources are used during perceptual processing. This assumption led the authors to focus on the role of the attentional modulation. In this account, if attention is directed to perceptual processing while the conceptual task is performed (and therefore occupying modality-specific attentional resources), it should interfere with processing of, for instance, a linguistic task of the same modality. Instead, if perceptual information is initially processed and no longer occupies attentional resources needed for its modality, leaving “that modality in an attentionally primed state” (Connell & Lynott, 2012; p. 4), a processing advantage should be observed in a conceptual task in the same modality.

In Chapter 4, we outlined a potential explanation of such findings based on Connell and Lynott’s account. Taking into account Chen and Mirman’s predictions and our discussion on the co-indexing mechanism, a potential explanation would be the following; first participants examine the visual context, two playing cards presenting written words.

Subsequently, they begin to read the critical sentence, such as '*Sandra met cheerfully her aunt at the health center*'. As they encounter the subject of the sentence ('*Sandra...*'), comprehenders' internal attention is direct to the representation of the first card from the visual context (since it presented the same word, "Sandra"). Participants continue reading the verbal phrase '*...met cheerfully...*' which theoretically should be co-indexed with the spatial distance between cards through spreading activation of the concept friendliness implied by the adverb. At this point it is possible that internal attention is directed extensively to the visual representation, thus producing a strong activation of the spatial representation, perhaps triggered by the fact that one aspect of the visual context (namely, the sentence object written in the second word) has not yet been encountered in the sentence.

This explanation would be in line with the two accounts briefly discussed here (Chen & Mirman, 2012; Connell & Lynott, 2012); as attention is steadily maintained in the spatial representation, which could be translated in a strong (rather than weak) activation, related conceptual representations are inhibited, resulting in longer reading times compared to unrelated conceptual representations. This explanation, however, comes with the caveat that results from Experiment 4 were not replicated in Experiment 5. Further investigation of this issue should shed light on whether this argument is substantial or whether the interference effect observed and reported here is accidental.

5.3.2. A roadmap for processing accounts of visually situated (abstract) language

The presented empirical findings posit a number of interesting challenges to existing accounts of situated language comprehension. Current models of situated sentence processing (Altmann & Kamide, 2007; Knoeferle & Crocker, 2007) have so far examined (and accounted for) a visual context that explicitly relates to critical linguistic stimuli. Generally, critical sentences are a description of the visual scene, which in turn is a depiction of the situation entailed by the sentence (or at least one state of it, in the case of a static visual context, see

Altmann & Kamide, 2007). Indeed, experiments in which an array of objects is presented in an imaginary grid (e.g., Duñabeitia et al. 2009; Huettig & Altmann, 2005) while participants hear a sentence, are not exact descriptions of the event described by the sentence. Such experiments, however, have shown effects that are relatively independent from sentence interpretation, and rather limited to word recognition (see Huettig & Altmann, 2007). In contrast, the experiments in this thesis demonstrated that visual contexts can modulate sentence interpretation, even when they are not an explicit depiction of the situation entailed in the sentence. Importantly, these effects are not tied to specific lexical-semantics, but instead emerge from incremental interpretation.

To this end, accounts of situated language processing would need to reformulate their assumptions with regards to the relationship between visual environments and language that enables these sources of information to interact rapidly. In this respect, theories of grounded cognition (Barsalou, 1999, Zwaan, 2004), among them conceptual metaphor theory (Lakoff & Johnson, 1980; Lakoff, 1987), have long since suggested that mapping between language and situated and embodied experience derived from the recurrent interactions in the real world, go beyond the literal description of objects and events. A central mechanism in conceptual metaphor theory, metaphorical mapping, identifies specific relations between perceptual representations and abstract language, which in combination with online mechanisms of language processing (such as co-indexing as we proposed) allow clear predictions about the critical visual information that can be relevant for language processing even in the non-referential visual context. Both models discussed here (Altmann & Kamide, 2007; Knoeferle & Crocker, 2007) advocate for an important role of mapping between language and scene in situated language. Metaphorical mapping is a subtler kind of mapping, which nevertheless seems to have consequences for language understanding.

An affordance-based account (on its current version, Altmann & Kamide, 2007), cannot accommodate this kind of subtle mapping. From a traditional viewpoint, affordances

are a range of possible actions that a viewer can infer from an object (see Gibson, 1977). Altmann and Kamide (2007) use a wider definition, suggesting that affordances “refer to knowledge based on our experience of how we interact with an object, and how that object interacts with other objects” (p. 510). The conflict with this viewpoint is that affordances, even in its extended definition, are still limited to actions that can be performed on, with or to objects. As discussed earlier in this thesis, abstract language does not refer to objects. Perhaps, in the real world, objects with similar characteristics are, usually physically close to each other. However, they most probably can still afford to be separated in space. Similarly, people that have intimate relations might in general be closer to one another. Yet, they may also afford to be physically distant. How about “peace and war”? These two concepts are undoubtedly related, they are antonyms. Yet, they cannot afford to be close together or far apart, at least not in a literal sense, since they are not physical entities. They might be, though, in a metaphorical sense. Consequently, an affordance-based account is not well equipped to accommodate visual context influences on abstract language as those from our studies.

Accordingly, the CIA (in its current version, Knoeferle & Crocker, 2007), is also insufficient to accommodate visually situated abstract language. It has so far proposed that mapping between language and visual scenes occur between nouns and objects or verbs and actions, through reference and co-indexing. Thus, the present findings require a refinement of the co-indexing mechanism in the CIA. In Knoeferle & Crocker (2007), co-indexing works as an integration mechanism, while attentional processes are mainly governed by language-mediated referential or anticipatory searching. Our results however, suggest that visual information and language can co-index even in a non-referential visual context, and, they tentatively suggest that linking objects to content words is not a condition of such interaction. Confronted with this evidence, it is plausible to argue that activation of long-term knowledge (stage 1 in the CIA) does not only activate expectations (for predictions), but also activates associations stored in semantic memory. This has a direct effect on language-mediated

attention (stage 2). In a non-referential environment referential/anticipatory search is not relevant (since there are no objects to be mentioned). Instead, spreading activation moves attention towards (or gives a boost of activation to) associated features of the mental representation of the visual context. This priming effect allows for co-indexing between language and visual information independently from referential links. If co-indexing does not require a referential link, it could be extended from explicit association (i.e., between nouns/verbs and objects/action) to subtler associations such as those proposed by conceptual metaphor theory.

Let's consider an example from Experiment 3, from a perspective in which spatial distance is co-indexed with semantic similarity. In Experiment 3, participants first examine the visual context, which depicted spatial distance between playing cards (i.e., either far apart or close together), and this information is stored in working memory (in our paradigm, due to the picture-verification task). Subsequently, participants begin to read a sentence, encountering the *and*-coordination of nouns '*Peace and war...*'. The second noun ('*war*'), embedded on its preceding linguistic context ('*Peace and...*') activates long-term representations (world-knowledge), as soon as it is first encountered. This means that the second noun implies similarity (or dissimilarity, in this case). This activation of the concept of similarity spreads rapidly directing attention (regardless of reference) to aspects of the visual context active in working memory that are associated with similarity. According to conceptual metaphor theory, the abstract concepts of similarity and the perceptual representation of distance are fundamentally related via metaphorical mapping. Thus, as similarity is implied by the *and*-coordination of nouns, it boosts activation of the spatial distance representation in working memory, making it available for current linguistic interpretation.

In this example, we can see that co-indexing still links linguistic meaning, visual information and world-knowledge, as in its original version. What this notion adds is an

explicit characterization of co-indexing, which does not require (but could work in the context of) a referential visual context, and more importantly that could accommodate a wider range of world-language relationships. This version of the co-indexing mechanism can accommodate findings in which direct lexical-semantic associations guide attention to unmentioned objects (e.g., Huettig & Altmann, 2005; Mirman & Magnuson, 2009), but also can accommodate subtler relations between visual and linguistic representations (e.g., Wassenburg & Zwaan, 2010; Coppens et al. 2012). Finally, it provides a mechanism that can explain how visual representations (including low-level spatial information) are integrated with real-time abstract language comprehension.

In summary, we proposed that metaphorical mapping creates the relevant links (perhaps during language acquisition, see Johnson, 1999) between concrete experiential representations and abstract concepts; during online language processing, spreading activation based on long-term association moves the attention of the cognitive system to related representations, boosting the activation of those (visual) representations active in working memory. Finally, different sources of information are co-indexed, allowing the integration of working memory contents, visual representations and world-knowledge.

5.4. Conclusions and future directions

This thesis presents the results of five eye-tracking reading experiments from which a number of inferences could be drawn. Most of these conclusions are closely related, however, a summary and an individualization of them could help to inspire future directions of research based on findings about a specific phenomenon of interest. This section is intended to provide a condensed version of the take-home message, so to speak. We separate this short summary in three general conclusions with regards to the relationship between the visual environment and language processing, and two further specific conclusions with regards to the mechanisms

discussed in the previous section. At the end of this section we present three new potential directions based on our findings, that we believe could deliver valuable information to extend and improve our understanding on situated language comprehension.

5.4.1. General conclusions

The first conclusion, and perhaps most challenging for the accounts of situated language comprehension discussed in this thesis, is that visual contexts do not need to literally depict objects, situations or any potential state thereof, subsequently described in a sentence, in order to influence the processing of that sentence incrementally. This finding contradicts a main assumption of existing models of sentence processing in context and demands they be updated. This finding extended previous results from visual-world studies, in which concrete and abstract words rapidly guide visual attention to unmentioned objects that are lexically related to them, to situations in which unrelated visual context affected language comprehension beyond lexical associations.

The second general conclusion is that abstract language can be rapidly and incrementally modulated by visual information at the compositional level. As discussed in Chapter 2, psycholinguistic research studying the visual world and language interactions has rarely examined visually situated abstract language comprehension. In contrast, theories of grounding cognition have argued that perceptual information (such as visual representations), are relevant for both abstract and concrete language comprehension. Yet, they have not extensively examined this question from an incremental language processing approach. In this context, the results our experiments also contribute to these theories, showing that as for concrete representation (in language), visual information plays an important role during online abstract sentence processing.

A third point is that our results showed that a low-level visual representation, namely spatial distance between objects, influences sentence reading times rapidly and incrementally. Previous research suggested that low-level visual representation could affect end-of-sentence response times (see, e.g., Zwaan et al., 2002), which did not allow for clear conclusions with regards to the time course of these effects. By using an eye-tracking methodology, we were able to show that spatial distance can affect abstract language interpretation in the early stages of sentence processing (first-pass reading times), and additionally in those sentence regions that critically expressed or implied the semantic content associated with spatial distance and beyond.

5.4.2. Specific conclusions

The first conclusion with regards to the mechanisms we suggest explain our results is that conceptual overlap and spreading activation between visual and linguistic representations (or the event structure implied by a sentence), can happen beyond affordances. Abstract concepts, such as similarity and intimacy, are related to spatial distance. However, this is not the same as to suggest that similarity affords closeness. In fact, similarity does not depend on spatial distance, and neither does intimacy. Similar objects can be located either far apart or close together; family members can live together or in different cities. This finding does not contradict an affordance-based account. We argue that a spreading activation mechanism can still account for the relationship between spatial distance and abstract concepts such as similarity and intimacy. However, it requires an extension that could explain visual world-language interactions that go beyond concrete objects and actions.

The second conclusion at this level concerns the co-indexing mechanism. In the version presented as part of CIA, co-indexing integrates visual, linguistic and world knowledge information for concrete objects and actions. Therefore, it cannot account for links

between spatial distance and abstract concepts such as similarity or intimacy. We propose, that this mechanism should be extended to subtler relationships beyond nouns/verbs and objects/actions, such as those between adjectives and spatial representations. For the CIA, this means that the establishment of references would not be condition for the interaction between visual context and language processing, extending its explanatory power to situations in which attention is guided to unmentioned objects. It would also require a more explicit notion of spreading activation from world-knowledge to (perhaps also subtler) aspects of the visual context that might be related.

5.4.3. New directions

These results not only presented a number of challenges to existing models of situated language processing. They also offer novel empirical findings that create a space for a new direction in the study of language in context. In this regard, the first footprint that the present findings leave is that of non-referential visual context and its influence on incremental language processing. This unique finding opens the opportunity to study a whole range of subtler relations between visual information and language understanding (perhaps even language production). It demands, however, a theory that can make clear predictions with regards to the kind of non-referential visual information that can be relevant to specific linguistic representations or processes. Conceptual metaphor theory, for instance, describes particularly distinctive connections between perceptual representations (some of them visual in nature) and abstract concepts, through metaphorical mapping. We believe, however, that this movement towards ever subtler relationships between visual information and language processing is not restricted to a conceptual metaphors framework. In fact, during language comprehension people often use visual information that is not explicitly mentioned in, or lexically-associated with, the utterance they are processing. For instance, emotional cues such

as facial expression and body language are indeed relevant for language processing (see, e.g., Carminati & Knoeferle, 2013; Etcoff, Ekman, Magee, & Frank, 2000), yet these sources of information have not extensively studied in the context of real-time situated language processing. The present work offers an account that recognizes the potential effects of non-referential visual context on incremental sentence comprehension.

The second important direction that this work could lead to is the study of visually situated abstract language. As discussed in Chapter 2, psycholinguistic research has concentrated mainly on examining how concrete language interacts with visual information from objects and actions. We have shown that processing of abstract nouns and abstract adjectives/adverbs embedded in a sentence context can also be affected by visual information. In this thesis we focused on two conceptual metaphors, namely SIMILARITY IS CLOSENESS and INTIMACY IS CLOSENESS, yet there are dozens of conceptual metaphors that have been described by Lakoff and colleagues (e.g., Lakoff & Johnson, 1999; Lakoff et al., 1991 for a compilation). The present findings strongly suggest that metaphorical mapping has consequences for online language processing. Consequently, another future direction that this research project could take is the examination of further relationships between perceptual and abstract representations.

One important issue for models of sentence processing, which this thesis did not directly address, is the notion of prediction during language comprehension. We discussed in Chapter 2 how the interaction between visual information, world-knowledge and language clearly allow participants to predict incoming linguistic information. Could spatial representations be used to draw inferences and predictions in similar ways? This is an empirical question that perhaps demands one step back to scenes that somehow describe potential states of situations that are verbally described, for instance, in a visual-world paradigm with spoken sentence. Even then, an interesting question is whether low-level visual

information (e.g., distance between objects), in conjunction with abstract language (e.g., similar, dissimilar) could allow comprehenders to predict the object that is more likely to be subsequently mentioned.

In summary, this research project could push the study of visually situated language comprehension in (at least) three potential new directions; firstly, it appeals for even subtler world-language connections than those so far studied; secondly, it calls for the inclusion of abstract language in studies and models within psycholinguistics and thirdly, it entertains the possibility that low-level representation and abstract language could be used during predictive language comprehension.

In summary, this research project could push the study of visually situated language comprehension in (at least) three potential new directions; firstly, it appeals for even subtler world-language connections than those so far studied; secondly, it calls for the inclusion of abstract language in studies and models within psycholinguistics and thirdly, it entertains the possibility that low-level representation and abstract language could be used during predictive language comprehension.

GERMAN SUMMARY

Experimentelle Beweise haben gezeigt, dass visuelle Informationen schnell und inkrementell mit Sprachverständnis interagieren, wenn Sätze, Objekte und Ereignisse eine Szene beschreiben (Knoeferle et al, 2005. Tanenhaus et al, 1995). Allerdings liegen fast keine Studien vor, die untersucht haben, ob Online-Satzverarbeitung und visuelle Umgebungen interagieren können, wenn die Sätze keine Objekte oder Aktionen in einer visuellen Umgebung beschreiben. Infolgedessen konzentriert sich diese Dissertation auf zwei Fragen, die im Zusammenhang mit der Forschung in Psycholinguistik bisher wenig erforscht sind. Zuerst stellen wir die Frage, ob abstrakte Sprache, die nicht direkt über Objekte und Ereignisse spricht, auch schnell und inkrementell mit visuellen Informationen zusammenwirken. Zweitens stellen wir die Frage, ob visuelle Kontexte, die keine Darstellung von Objekten und Ereignissen in einem Satz beschreiben, die Lesegeschwindigkeit für diesen Satz beeinflussen können. Darüber hinaus soll diese Arbeit helfen, diejenigen kognitiven Mechanismen zu erklären, die eine schnelle Interaktion zwischen abstrakter Sprache und dem, was wir als *nicht-referenzielle* visuelle Zusammenhänge bezeichnen, ermöglichen.

Zur Beantwortung dieser Fragen haben wir im ersten Schritt die vorhandene Literatur über die Interaktion visueller Welt-Sprache detailliert untersucht. Im folgenden Kapitel werden eine Reihe von psycholinguistischen Studien betrachtet sowie ein breites Spektrum von Fragestellungen in Bezug auf visuelle Satzverarbeitung untersucht. Diese Überprüfung

zeigte eine Tendenz auf, von klaren und expliziten Zusammenhängen zwischen dem Gesehenen und der Bedeutung des Satzes, zu subtilen Beziehungen zwischen visuellen Kontexten und Sprachbedeutung (z.B. Wortassoziationen). Wir zogen dann zur Sichtung Studien heran, die nicht direkt im Zusammenhang mit Online-Satzverarbeitung standen, aber welche die Beziehung zwischen visuell-räumlichen Darstellungen und Sprachverständnis untersuchten. Wir konnten feststellen, dass sich die meisten Studien auf konkrete Sprache fokussieren und die Ergebnisse in Bezug auf abstrakte Sprache nicht eindeutig oder auf Worterkennung beschränkt sind. Im Gegensatz dazu gibt es Studien, die sich tatsächlich auf abstrakte Konzepte konzentrieren. Diese Studien deckten eine Reihe von verschiedenen konzeptionellen Aufgaben auf, wie Entscheidungen über Ähnlichkeit, schnelle Kategorisierung und Bewertung. Sie erforschten mehrere abstrakte Konzepte, darunter die Begriffe der Ähnlichkeit, Zeit, Intimität, Kategorien und affektiver Wertigkeiten, sowie die konkreten Erfahrungen, die mit ihnen verbunden sind, wie beispielsweise räumliche Distanz, Rückhaltung und räumlichen Positionen. Dennoch zeigte keine dieser Studien Beweise für die Beteiligung solcher visueller Darstellungen für das Online-Sprachverständnis.

Von diesen abstrakten Konzepten konzentrieren wir uns erstens auf die Konzepte der Ähnlichkeit und der Intimität, für die wir überzeugende theoretische und empirische Hinweise gefunden haben, die eine enge Verbindung zu einer bestimmten räumlichen Darstellung, nämlich räumlicher Distanz (nah beieinander vs. weit auseinander) andeuten. Diese zwei abstrakten Konzepte sind durch eine linguistische Theorie als konzeptuelle Metaphern beschrieben (Lakoff & Johnson, 1999), da ihre Verankerung in der Alltagssprache zu beobachten ist. Wir diskutierten über vorhandene Verhaltenshinweise für diese beiden konzeptuellen Metaphern und argumentieren, dass sie geeignet sind, um die Beziehung zwischen räumlichen Darstellungen und Online-Verarbeitung abstrakter Sätze zu untersuchen. Am Ende des Kapitels stellen wir zwei vorhandene Modelle der Satzverarbeitung in visuellem Kontext vor (Altmann & Kamide, 2007; Knoeferle & Crocker, 2007). Wir diskutieren über

ihre zentralen kognitiven Mechanismen und evaluieren wie sie potentielle nicht-referenzielle Auswirkungen auf Online-Satzverarbeitung und die Interaktion zwischen visueller Information und abstrakter Satz-Interpretation erklären. Wir argumentieren, dass diese Modelle in der aktuellen Fassung noch nicht ausreichend geeignet sind, um zu erklären, ob visuell gelegenes abstraktes Sprachverständnis oder visuelle Szenen, die keine Objekte oder Ereignisse verbal im Satz beschreiben, eine potentielle Gefährdung der Verarbeitung dieses Satzes bedeuten.

In den Kapiteln 3 und 4 präsentierten wir ein neues experimentelles Eye-Tracking-Paradigma, das entworfen wurde, um die beiden Fragestellungen in Kapitel 1 zu untersuchen. In fünf Experimenten wendeten wir zwei Ebenen der räumlichen Distanz mit zwei Ebenen des abstrakten semantischen Inhaltes an, sodass jedes Experiment vier experimentelle Bedingungen hatte. Für die Teilnehmer wurde ein visueller Kontext vorgestellt, in dem zwei Spielkarten sich entweder nahe beieinander oder weit auseinander (für experimentelle Durchgänge) bewegten. Unmittelbar danach lasen die Teilnehmer einen Satz, der semantische Ähnlichkeit zwischen zwei abstrakten Substantiven (Experiment 1, 2 und 3) oder einer sozialen Interaktion zwischen zwei Zeichen (Experiment 4 und 5). Die Augenbewegungen wurden aufgenommen während die Teilnehmer die Sätze gelesen haben. Im Anschluss wurden drei Messungen berechnet. Die wesentlichen Bereiche des Satzes waren diejenigen, in denen Ähnlichkeit (entweder ähnlich oder unähnlich) oder die Art der sozialen Beziehung (entweder freundlich oder unfreundlich) implizit oder explizit erwähnt wurden.

Die Ergebnisse der ersten drei Experimente zeigten signifikante Interaktionen zwischen räumlicher Distanz und semantischer Ähnlichkeit zwischen zwei abstrakten Substantiven; Sätze, die Ähnlichkeit ausdrücken, wurden schneller gelesen, wenn sich die Spielkarten im visuellen Kontext zueinander bewegten als wenn die Spielkarten sich voneinander entfernten. Im Gegensatz dazu wurden Sätze, die Unähnlichkeit ausdrücken,

schneller gelesen, wenn sich die Spielkarten im visuellen Kontext voneinander entfernten als wenn die Spielkarten sich zueinander bewegten. Soweit wir wissen, sind das die ersten Beweise dafür, dass nicht-referenzielle visuelle Kontexte abstraktes Sprachverständnis schnell und inkrementell modulieren. Die Analyse der Eye-Tracking-Daten aus Experiment 4 erweitern diese Ergebnisse und zeigen, dass die räumliche Distanz auch Sätze über soziale Beziehungen modulieren kann. Interessanterweise zeigten diese Ergebnisse das gegenteilige Muster der Experimente über Ähnlichkeit; Sätze, die eine freundliche soziale Interaktion zwischen zwei Zeichen ausdrücken, wurden schneller gelesen, wenn die Spielkarten in dem visuellen Kontext sich voneinander entfernten, im Vergleich zu den Sätzen, in denen die Spielkarten sich zueinander bewegten und umgekehrt für Sätze, die eine unfreundliche soziale Interaktion ausdrücken. Allerdings wurden diese Effekte nicht in einem nachfolgenden Experiment repliziert, was darauf hindeutet, dass vielleicht vollständig nicht-referenzieller visueller Kontext abstrakte Sprache nur unter bestimmten Umständen modulieren könnte.

In Kapitel 5 diskutieren wir unsere Literaturrecherche, welche die wichtigsten Ergebnisse in der Erforschung des visuell gelegen Sprachverständnisses zusammenfasst und aufzeigt, wie wir kritische offene Fragen identifizieren und sie anschließend als Forschungsfragen formulieren. Dann konzentrieren wir uns auf die Diskussion, inwieweit unsere Ergebnisse unsere Forschungsfragen beantworten können. Anschließend haben wir unsere Ergebnisse mit Annahmen und Prognosen aus bestehenden Modellen von Sprachverständnis in visuellem Kontext gegenübergestellt und diskutierten über einige der Herausforderungen, die diese Erkenntnisse bieten. Wir argumentierten, dass diese Modelle in ihrem gegenwärtigen Zustand nicht fähig sind unsere Erkenntnisse aufzunehmen. Sie deuten aber an, dass eine Erweiterung der bestehenden Mechanismen wie Ko-Indexieren und Aktivierungsausbreitung unsere Daten erklären könnten. Folglich haben wir beschrieben, wie diese aktualisierten Mechanismen die Auswirkungen der nicht-referenziellen visuellen Zusammenhänge auf Online-Verständnis abstrakter Sprache erklären könnten. Am Ende des

Kapitels schließen wir mit einer Zusammenfassung der wichtigsten Schulssfolgerungen und einigen Anmerkungen über zukünftige Entwicklungen und Herausforderungen für dieses Forschungsprogramm.

APPENDIX: Item Sentences

Critical sentences in experiments 1, 2 and 3

Forty-eight experimental sentences, each with two versions; either expressing dissimilarity (A.) or similarity (B.) between two abstract nouns.

Item 1

- A. Souveränität und Unsicherheit sind geradezu umgekehrt, so argumentierte die Denkerin.
- B. Verlegenheit und Unsicherheit sind ziemlich einheitlich, so argumentierte die Denkerin.

Item 2

- A. Munterkeit und Mattigkeit sind geradezu umgekehrt, so begründete die Logopädin.
- B. Ermüdung und Mattigkeit sind ziemlich einheitlich, so begründete die Logopädin.

Item 3

- A. Hochverrat und Loyalität sind geradezu umgekehrt, das behauptete die Fußballspielerin.
- B. Ergebenheit und Loyalität sind ziemlich einheitlich, das behauptete die Fußballspielerin.

Item 4

- A. Fortuna und Unglück sind sicher ungleich, so ersann der Clown.
- B. Unheil und Unglück sind gewiss verwandt, so ersann der Clown.

Item 5

- A. Ignoranz und Vorgefühl sind bestimmt verschieden, so träumte das Medium.
- B. Vorahnung und Vorgefühl sind freilich entsprechend, so träumte das Medium.

Item 6

- A. Frieden und Krieg sind bestimmt verschieden, das verriet der Anthropologe.
- B. Kampf und Krieg sind freilich entsprechend, das verriet der Anthropologe.

Item 7

- A. Werktag und Ruhepause sind sicher ungleich, das beschwur der Arbeiter.
- B. Atempause und Ruhepause sind gewiss verwandt, das beschwur der Arbeiter.

Item 8

- A. Ohnmacht und Können sind bestimmt verschieden, das verkündete der Artist.
- B. Talent und Können sind freilich entsprechend, das verkündete der Artist.

Item 9

- A. Freudenruf und Wutanfall sind sicher ungleich, so bekundete der Boxer.
- B. Zornausbruch und Wutanfall sind gewiss verwandt, so bekundete der Boxer.

Item 10

- A. Betrug und Benehmen sind eher andersartig, so verkündigte der Polizeibeamte.
- B. Anstand und Benehmen sind fast äquivalent, so verkündigte der Polizeibeamte.

Item 11

- A. Dummheit und Weisheit sind bestimmt verschieden, das erklärte der Professor.
- B. Begabung und Weisheit sind freilich entsprechend, das erklärte der Professor.

Item 12

- A. Lebensfreude und Traurigkeit sind geradezu umgekehrt, so tönte die Dichterin.
- B. Melancholie und Traurigkeit sind ziemlich einheitlich, so tönte die Dichterin.

Item 13

- A. Frohsinn und Trübsinn sind geradezu umgekehrt, so empfahl die Frau.
- B. Schwermut und Trübsinn sind ziemlich einheitlich, so empfahl die Frau.

Item 14

- A. Diebstahl und Bezahlung sind geradezu umgekehrt, so schrie die Gewerkschaftlerin.
- B. Vergütung und Bezahlung sind ziemlich einheitlich, so schrie die Gewerkschaftlerin.

Item 15

- A. Verstimmung und Wohlbefinden sind eher andersartig, so ließ die Kinderärztin verlauten.
- B. Zufriedenheit und Wohlbefinden sind fast äquivalent, so ließ die Kinderärztin verlauten.

Item 16

- A. Infektion und Gesundheit sind geradezu umgekehrt, so dachte die Krankenschwester.
- B. Genesung und Gesundheit sind ziemlich einheitlich, so dachte die Krankenschwester.

Item 17

- A. Betrübnis und Euphorie sind eher andersartig, so antwortete die Lehrerin.
- B. Entzücken und Euphorie sind fast äquivalent, so antwortete die Lehrerin.

Item 18

- A. Trennung und Heirat sind geradezu umgekehrt, so beurteilte die Organisatorin.
- B. Hochzeit und Heirat sind ziemlich einheitlich, so beurteilte die Organisatorin.

Item 19

- A. Minorität und Majorität sind eher andersartig, das erkannte die Präsidentin.
- B. Übermacht und Majorität sind fast äquivalent, das erkannte die Präsidentin.

Item 20

- A. Unwohlsein und Wohlgefühl sind eher andersartig, so erwiderte die Psychologin.
- B. Wohlbehagen und Wohlgefühl sind fast äquivalent, so erwiderte die Psychologin.

Item 21

- A. Bestrafung und Belohnung sind eher andersartig, das bestätigte die Richterin.
- B. Besoldung und Belohnung sind fast äquivalent, das bestätigte die Richterin.

Item 22

- A. Kräftigkeit und Schwachheit sind eher andersartig, das beteuerte die Sportlehrerin.
- B. Ermattung und Schwachheit sind fast äquivalent, das beteuerte die Sportlehrerin.

Item 23

- A. Aufregung und Entspannung sind eher andersartig, das bekräftigte die Therapeutin.
- B. Erholung und Entspannung sind fast äquivalent, das bekräftigte die Therapeutin.

Item 24

- A. Feindschaft und Kameradschaft sind eher andersartig, das enthüllte die Beraterin.
- B. Partnerschaft und Kameradschaft sind fast äquivalent, das enthüllte die Beraterin.

Item 25

- A. Zustimmung und Widerspruch sind bestimmt verschieden, so entschied der Lehrer.
- B. Opposition und Widerspruch sind freilich entsprechend, so entschied der Lehrer.

Item 26

- A. Illusion und Gewissheit sind bestimmt verschieden, das erzählte der Wissenschaftler.
- B. Tatbestand und Gewissheit sind freilich entsprechend, das erzählte der Wissenschaftler.

Item 27

- A. Leid und Glanz sind eher andersartig, das äußerte das Fotomodell.
- B. Ruhm und Glanz sind fast äquivalent, das äußerte das Fotomodell.

Item 28

- A. Abneigung und Zuneigung sind sicher ungleich, so meinte der Großvater.
- B. Zuwendung und Zuneigung sind gewiss verwandt, so meinte der Großvater.

Item 29

- A. Konflikt und Einigung sind bestimmt verschieden, das fand der Historiker.
- B. Eintracht und Einigung sind freilich entsprechend, das fand der Historiker.

Item 30

- A. Rückzug und Offensive sind sicher ungleich, so brüllte der Infanterist.
- B. Feldzug und Offensive sind gewiss verwandt, so brüllte der Infanterist.

Item 31

- A. Faktum und Märchen sind bestimmt verschieden, so las der Journalist.
- B. Fiktion und Märchen sind freilich entsprechend, so las der Journalist.

Item 32

- A. Unbegabtheit und Kreativität sind geradezu umgekehrt, so korrigierte die Wahrsagerin.
- B. Genialität und Kreativität sind ziemlich einheitlich, so korrigierte die Wahrsagerin.

Item 33

- A. Humor und Trauer sind eher andersartig, so deklarierte das Mädchen.
- B. Kummer und Trauer sind fast äquivalent, so deklarierte das Mädchen.

Item 34

- A. Jammer und Pracht sind bestimmt verschieden, das glaubte der Mann.
- B. Luxus und Pracht sind freilich entsprechend, das glaubte der Mann.

Item 35

- A. Gewinn und Verlust sind sicher ungleich, so zeterte der Politiker.
- B. Schaden und Verlust sind gewiss verwandt, so zeterte der Politiker.

Item 36

- A. Hochgefühl und Höllenangst sind bestimmt verschieden, so rezitierte der Poet.
- B. Herzensangst und Höllenangst sind freilich entsprechend, so rezitierte der Poet.

Item 37

- A. Friedsamkeit und Zwietracht sind geradezu umgekehrt, das rief die Priesterin.
- B. Streiterei und Zwietracht sind ziemlich einheitlich, das rief die Priesterin.

Item 38

- A. Übereinkunft und Kampfhandlung sind bestimmt verschieden, das sagte der Soldat.
- B. Konfrontation und Kampfhandlung sind freilich entsprechend, das sagte der Soldat.

Item 39

- A. Dissonanz und Wohlklang sind sicher ungleich, das bejahte der Sänger.
- B. Wohllaut und Wohlklang sind gewiss verwandt, das bejahte der Sänger.

Item 40

- A. Glück und Sorge sind sicher ungleich, das schrieb der Schauspieler.
- B. Angst und Sorge sind gewiss verwandt, das schrieb der Schauspieler.

Item 41

- A. Fairness und Mogelei sind sicher ungleich, das weissagte der Schiedsrichter.
- B. Gaunerei und Mogelei sind gewiss verwandt, das weissagte der Schiedsrichter.

Item 42

- A. Antipathie und Liebelei sind eher andersartig, das stellte der Schreiber fest.
- B. Romanze und Liebelei sind fast äquivalent, das stellte der Schreiber fest.

Item 43

- A. Arbeitszeit und Vergnügen sind sicher ungleich, das suggerierte der Unterhalter.
- B. Unterhaltung und Vergnügen sind gewiss verwandt, das suggerierte der Unterhalter.

Item 44

- A. Verachtung und Verehrung sind bestimmt verschieden, so vermutete die Dame.
- B. Ehrfurcht und Verehrung sind freilich entsprechend, so vermutete die Dame.

Item 45

- A. Belobigung und Entwürdigung sind geradezu umgekehrt, das verteidigte die Soziologin.
- B. Erniedrigung und Entwürdigung sind ziemlich einheitlich, das verteidigte die Soziologin.

Item 46

- A. Vergötterung und Nichtbeachtung sind geradezu umgekehrt, das zeigte die Nonne.
- B. Geringachtung und Nichtbeachtung sind ziemlich einheitlich, das zeigte die Nonne.

Item 47

- A. Achtsamkeit und Zerstreuung sind sicher ungleich, das teilte der Psychiater mit.
- B. Zerfahrenheit und Zerstreuung sind gewiss verwandt, das teilte der Psychiater mit.

Item 48

- A. Mittellosigkeit und Besitzlosigkeit sind sicher ungleich, so berichtigte der Wirtschaftler.
- B. Wohlhabenheit und Besitzlosigkeit sind gewiss verwandt, so berichtigte der Wirtschaftler.

Critical sentences in experiments 4 and 5

Forty-eight experimental sentences, each with two versions; either expressing an unfriendly (A.) or a friendly (B.) interaction between two characters.

Item 1

- A. Alexander grüßt seinen Onkel grob auf dem Gehweg.
- B. Alexander grüßt seinen Onkel sanft auf dem Gehweg.

Item 2

- A. Bastian verabschiedet seinen Großvater grob vor dem Kino.
- B. Bastian verabschiedet seinen Großvater sanft vor dem Kino.

Item 3

- A. Benjamin begegnet seiner Schwägerin grob auf der Familienfeier.
- B. Benjamin begegnet seiner Schwägerin sanft auf der Familienfeier.

Item 4

- A. Daniel informiert seinen Neffen unfreundlich über die Urlaubsplanung.
- B. Daniel informiert seinen Neffen wohlwollend über die Urlaubsplanung.

Item 5

- A. Hannah beratschlagt ihre Patentante unfreundlich nach dem Treffen.
- B. Hannah beratschlagt ihre Patentante wohlwollend nach dem Treffen.

Item 6

- A. Jürgen benachrichtigt seinen Stiefbruder rücksichtsvoll über den Befund.
- B. Jürgen benachrichtigt seinen Stiefbruder unbarmherzig über den Befund.

Item 7

- A. Karl betrachtet seinen Enkel lieblos während des Gesprächs.
- B. Karl betrachtet seinen Enkel warmherzig während des Gesprächs.

Item 8

- A. Katharina behandelt ihre Großtante rücksichtsvoll bei der Begrüßung.
- B. Katharina behandelt ihre Großtante unbarmherzig bei der Begrüßung.

Item 9

- A. Lena berät ihren Bruder respektvoll beim Abendessen.
- B. Lena berät ihren Bruder ungehalten beim Abendessen.

Item 10

- A. Leona berührt ihre Nichte lieblos beim Abschied.
- B. Leona berührt ihre Nichte warmherzig beim Abschied.

Item 11

- A. Linda beurteilt ihre Cousine unfreundlich in der Rede.
- B. Linda beurteilt ihre Cousine wohlwollend in der Rede.

Item 12

- A. Maike fragt ihren Großonkel gutgelaunt nach dem Wohlbefinden.
- B. Maike fragt ihren Großonkel misstrauisch nach dem Wohlbefinden.

Item 13

- A. Maria begrüßt ihren Urgroßvater lieblos vor der Haustür.
- B. Maria begrüßt ihren Urgroßvater warmherzig vor der Haustür.

Item 14

- A. Markus befragt seine Schwester rücksichtsvoll vor dem Termin.
- B. Markus befragt seine Schwester unbarmherzig vor dem Termin.

Item 15

- A. Melanie verständigt ihre Urgroßmutter respektvoll über die Erbschaft.
- B. Melanie verständigt ihre Urgroßmutter ungehalten über die Erbschaft.

Item 16

- A. Michaela schreibt ihrer Großmutter respektvoll während der Festtage.
- B. Michaela schreibt ihrer Großmutter ungehalten während der Festtage.

Item 17

- A. Oskar umarmt seinen Patenonkel respektvoll bei dem Ausflug.
- B. Oskar umarmt seinen Patenonkel ungehalten bei dem Ausflug.

Item 18

- A. Paul belehrt seine Ehefrau unfreundlich in den Briefen.
- B. Paul belehrt seine Ehefrau wohlwollend in den Briefen.

Item 19

- A. Sabine begleitet ihren Vater rücksichtsvoll zum Einkaufen.
- B. Sabine begleitet ihren Vater unbarmherzig zum Einkaufen.

Item 20

- A. Sandra trifft ihre Tante gutgelaunt in der Praxis.
- B. Sandra trifft ihre Tante missmutig in der Praxis.

Item 21

- A. Stefan mailt seiner Mutter gutgelaunt über die Geburtstagsparty.
- B. Stefan mailt seiner Mutter missmutig über die Geburtstagsparty.

Item 22

- A. Susanne erlebt ihren Opa grob bei der Umarmung.
- B. Susanne erlebt ihren Opa sanft bei der Umarmung.

Item 23

- A. Tina kontaktiert ihren Cousin gutgelaunt aus dem Urlaub.
- B. Tina kontaktiert ihren Cousin missmutig aus dem Urlaub.

Item 24

- A. Tom empfängt seine Freundin lieblos nach dem Wochenende.
- B. Tom empfängt seine Freundin warmherzig nach dem Wochenende.

Item 25

- A. Der Architekt benachrichtigt den Makler rücksichtsvoll über die Baumängel.
- B. Der Architekt benachrichtigt den Makler unbarmherzig über die Baumängel.

Item 26

- A. Der Bankier betrachtet den Börsianer lieblos beim Dinner.
- B. Der Bankier betrachtet den Börsianer warmherzig beim Dinner.

Item 27

- A. Der Diplomat verabschiedet den Politiker grob im Bundestag.
- B. Der Diplomat verabschiedet den Politiker sanft im Bundestag.

Item 28

- A. Der Finanzbeamte berät den Steuerberater respektvoll beim Beratungsgespräch.
- B. Der Finanzbeamte berät den Steuerberater ungehalten beim Beratungsgespräch.

Item 29

- A. Der Kapitän fragt die Reederin gutgelaunt nach der Regatta.
- B. Der Kapitän fragt die Reederin missmutig nach der Regatta.

Item 30

- A. Der Kellner verständigt die Barkeeperin respektvoll vor dem Event.
- B. Der Kellner verständigt die Barkeeperin ungehalten vor dem Event.

Item 31

- A. Der Kosmetiker schreibt dem Friseur respektvoll nach der Abschlussprüfung.
- B. Der Kosmetiker schreibt dem Friseur ungehalten nach der Abschlussprüfung.

Item 32

- A. Der Manager trifft die Prominente gutgelaunt zum Frühstück.
- B. Der Manager trifft die Prominente missmutig zum Frühstück.

Item 33

- A. Der Oberarzt berührt den Chirurgen lieblos bei der Visite.
- B. Der Oberarzt berührt den Chirurgen warmherzig bei der Visite.

Item 34

- A. Der Opernsänger begrüßt die Schauspielerin lieblos an der Garderobe.
- B. Der Opernsänger begrüßt die Schauspielerin warmherzig an der Garderobe.

Item 35

- A. Der Pilot behandelt die Fluglotsin rücksichtsvoll bei der Fortbildung.
- B. Der Pilot behandelt die Fluglotsin unbarmherzig bei der Fortbildung.

Item 36

- A. Der Verkäufer informiert den Kassierer unfreundlich über die Abrechnung.
- B. Der Verkäufer informiert den Kassierer wohlwollend über die Abrechnung.

Item 37

- A. Die Empfangsdame mailt dem Nachtwächter gutgelaunt nach Silvester.
- B. Die Empfangsdame mailt dem Nachtwächter missmutig nach Silvester.

Item 38

- A. Die Floristin begrüßt den Studenten grob im Laden.
- B. Die Floristin begrüßt den Studenten sanft im Laden.

Item 39

- A. Die Gräfin erlebt die Adelige grob während des Essens.
- B. Die Gräfin erlebt die Adelige sanft während des Essens.

Item 40

- A. Die Konditorin begegnet dem Zulieferer grob bei der Frühschicht.
- B. Die Konditorin begegnet dem Zulieferer sanft bei der Frühschicht.

Item 41

- A. Die Lehrerin befragt den Beamten rücksichtsvoll nach dem Elternsprechtag.
- B. Die Lehrerin befragt den Beamten unbarmherzig nach dem Elternsprechtag.

Item 42

- A. Die Mechanikerin beratschlagt die Installateurin unfreundlich nach dem Auftrag.
- B. Die Mechanikerin beratschlagt die Installateurin wohlwollend nach dem Auftrag.

Item 43

- A. Die Premierministerin begleitet die Präsidentin rücksichtsvoll zum Wagen.
- B. Die Premierministerin begleitet die Präsidentin unbarmherzig zum Wagen.

Item 44

- A. Die Professorin beurteilt den Dozenten unfreundlich im Treppenhaus.
- B. Die Professorin beurteilt den Dozenten wohlwollend im Treppenhaus.

Item 45

- A. Die Regisseurin empfängt den Produzenten lieblos bei der Preisverleihung.
- B. Die Regisseurin empfängt den Produzenten warmherzig bei der Preisverleihung.

Item 46

- A. Die Schaffnerin umarmt die Zugführerin respektvoll bei der Zugfahrt.
- B. Die Schaffnerin umarmt die Zugführerin ungehalten bei der Zugfahrt.

Item 47

- A. Die Staatsanwältin kontaktiert den Richter gutgelaunt vor Feierabend.
- B. Die Staatsanwältin kontaktiert den Richter missmutig vor Feierabend.

Item 48

- A. Die Vorsitzende belehrt die Direktorin unfreundlich vor dem Meeting.
- B. Die Vorsitzende belehrt die Direktorin wohlwollend vor dem Meeting.

REFERENCES

- Allopenna, P., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye-movements: evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419-439.
- Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, *93*, B79-B87.
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247-264.
- Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502–518.
- Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition*, *111*, 55–71.
- Altmann, G. T. M., & Mirković, J. (2009). Incrementality and Prediction in Human Sentence Processing. *Cognitive Science*, *33*, 583-609.
- Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191-238.
- Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, *31*, 685-712.
- Anderson, J. R., & Pirolli, P. L. (1984). Spread of activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 791-798.

- Baayen, R. H. (2008). *Analyzing Linguistic Data. A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.
- Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
- Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: evidence from a stroop task. *Journal of Experimental Psychology*, *136*, 610-622.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-660.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.
- Barsalou, L. W., Santos, A., Simmons, W. K. & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. A. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford, England: Oxford University Press.
- Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, *109*, 18–40.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.
- Breaux, B. O., & Feist, M. I. (2008). The color of similarity. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*. (pp. 253-258).

- Bergen, B. K., Lindsay, S., Matlock, T., Narayanan, S. (2007). Spatial and Linguistic Aspects of Visual Imagery in Sentence Comprehension. *Cognitive Science*, 31, 733-764.
- Bergen, B. K. (2007). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittleberg, S. Coulson, & M. Spivey (Eds.). *Methods in cognitive linguistics*. Philadelphia, PA: John Benjamins.
- Boot, I., & Pecher, D. (2010). Similarity is closeness: Metaphorical mapping in a conceptual task. *The Quarterly Journal of Experimental Psychology*, 63, 942-954.
- Boot, I., & Pecher, D. (2011). Representation of Categories Metaphorical Use of the Container Schema. *Experimental Psychology*, 58, 162–170.
- Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, 75, 1-28.
- Campana, E., Silverman, L., Tanenhaus, M. K., Bennetto, L., & Packard, S. (2005). Real-time integration of gesture and speech during reference resolution. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*. (pp. 378-383).
- Casasanto, D. (2008). Similarity and proximity: When does close in space mean close in mind? *Memory & Cognition*, 36, 1047-1056.
- Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, 106, 579-593.
- Carminati, M. N., & Knoeferle, P. (2013). Effects of speaker emotional facial expression and listener age on incremental sentence processing. *PloS one*, 8, 1-16.
- Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition*, 14, 509-522.

- Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing Referential Domains during Real-Time Language Comprehension. *Journal of Memory and Language*, *47*, 30-49.
- Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and Affordances in Syntactic Ambiguity Resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 687-696.
- Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*, 417-430.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.
- Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476-485.
- Connell, L., & Lynott, D. (2012). When does perception facilitate or interfere with conceptual processing? The effect of attentional modulation. *Frontiers in Psychology*, *3*, 1-4.
- 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.
- Coppens, L. C., Gootjes, L., Zwaan, R. A. (2012). Incidental picture exposure affects later reading: Evidence from the N400. *Brain & Language*, *122*, 64-69.

- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163-201
- Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*, 615–627.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*, 507-534.
- Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken- word recognition. *Psychonomic Bulletin & Review*, *12*, 453– 459.
- Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*, 284-292.
- Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000). Lie detection and language comprehension. *Nature*, *405*, 139-139.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348-368.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259-281.
- Gallese, V., & Lakoff, G. (2005). The Brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*, 455–479.

- Gibbs, R. W., & Colston, H. L. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, 6-4, 347-378.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving acting and knowing: Towards an ecological psychology*. Hillsdale, NJ: Erlbaum.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Cognitive Science*, 1, 586-596.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558-565.
- Glenberg, A. M., & Robertson, D. A. (1999). Indexical Understanding of Instructions. *Discourse Processes*, 28, 1-26.
- Grady, J. E. (1997). Theories are buildings revisited. *Cognitive Linguistics*, 8, 267-290.
- Grady, J. (2005). Primary metaphors as inputs to conceptual integration. *Journal of Pragmatics*, 37, 1595-1614.
- Grady, J., & Johnson, C. (1997). Converging Evidence for the Notions of subscene and primary scene. In *Proceedings of the 23th Annual Meeting of the Berkeley Linguistics Society*. (pp. 123-136).
- Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society*, 362, 801-811.

- Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, *57*, 596–615.
- Hebb, D. (1949). *The organization of behavior: A neuropsychological theory*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Henderson, J., & Ferreira, F. (Eds.) (2004). *Interface of language, vision and action: Eye movements and the visual world*. New York: Psychology Press.
- Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition*, *108*, 533-542.
- Huetig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23-B32.
- Huetig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, *15*, 985-1018.
- Huetig, F., & Altmann, G. T. M. (2011). Looking at anything that is green when hearing “frog”: How object surface colour and stored object colour knowledge influence language-mediated overt attention. *The Quarterly Journal of Experimental Psychology*, *64*, 122-145.
- Huetig, F., & Hartsuiker, R. J. (2008). When you name the pizza you look at the coin and the bread: Eye movements reveal semantic activation during word production. *Memory & Cognition*, *36*, 341-360.
- Huetig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*, 460-482.

- Huetting, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica, 121*, 65-80.
- Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*, 785-813.
- Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of verbal learning and verbal behavior, 22*, 485-508.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 104*, 130-136.
- Johnson, C. (1999). Metaphor vs. conflation in the acquisition of polysemy: The case of see. In M. Hiraga, C. Sinha, & S. Wilcox (Eds.). *Amsterdam studies in the theory and history of linguistic science series 4*. Philadelphia, PA: John Benjamins
- Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*, 37-55.
- Kaschak, P. M., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94*, B79-B89.
- Knoeferle, P., Crocker, M. W., Scheepers, C., Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition, 95*, 95-127.
- Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science, 30*, 481-529.

- Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, *57*, 519-543.
- Knoeferle, P., & Kreysa, H. (2012). Can speaker gaze modulate syntactic structuring and thematic role assignment during spoken sentence comprehension? *Frontiers in Psychology*, *3*, 1-15.
- Knoeferle, P., Carminati, M. N., Abashidze, D., & Essig, K. (2011). Preferential inspection of recent real-world events over future events: evidence from eye tracking during spoken sentence comprehension. *Frontiers in Psychology*, *1*, 1-12.
- Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. *Psychophysiology*, *48*, 495–506.
- Konieczny, L., Hemforth, B., Scheepers, C. & Strube, G. (1997). The role of lexical heads in parsing: Evidence from German. *Language and Cognitive Processes*, *12*, 307–348.
- Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. D. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology*, *140*, 14-34.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press
- Lakoff, G. (2008). The neural theory of metaphor. In R. Gibbs (Ed.). *The Cambridge handbook of metaphor and thought*. Cambridge, UK: Cambridge University Press.

- Lakoff, G. (2012). Explaining embodied cognition results. *Topics in Cognitive Science*, 4, 773-785.
- Lakoff, G., Espenson, J., & Schwartz, A. (1991). *The master metaphor list*. Technical Report, University of California at Berkeley.
- Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4, 195-208.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.
- Liversedge, S. P., Paterson, K.B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. Oxford, UK: Elsevier.
- Matthews, J. L., & Matlock, T. (2011). Understanding the link between spatial distance and social distance. *Social Psychology*, 42, 185-192.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4, 310-322.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88, 375-407.
- McGlone, M. S., & Harding, J. L. (1998). Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1211-1223.

- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*, 547-559.
- Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: associations between affect and vertical position. *Psychological Science, 15*, 243-247.
- Meier, B.P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white: automatic inferences about stimulus valence based on brightness. *Psychological Science, 15*, 82-87.
- Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition, 37*, 1026-1039.
- Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 92-103.
- Myung, J. Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition, 98*, 223-243.
- Núñez, R. E., Motz, B. A., & Teuscher, U. (2006). Time after time: The psychological reality of the ego-and time-reference-point distinction in metaphorical construals of time. *Metaphor and Symbol, 21*, 133-146.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.
- Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York, NY: Cambridge University Press.

- Pecher, D., Boot, I., & van Dantzig, S. (2011). Abstract concepts: Sensory-motor grounding, metaphors, and beyond. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, 54 (pp. 217-248). Burlington: Academic Press.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633-651.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, 134, 427-459.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576-582.
- Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 17, 458-470.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11, 351-360.
- Quené, H. & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43, 103-121.
- Quené, H. & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59, 413-425.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K. & Liveriedge, S. P. (2004). Visual and linguistic processing during eye fixation in reading. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action. Eye movements and the visual world*. New York, NY: Psychology Press.

- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: looking at things that aren't there anymore. *Cognition*, *76*, 269-295.
- Richardson, D. C., Spivey, M. J., Barsalou, L. W., McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*, 767-780.
- Richardson, D. C., Spivey, M. J., Edelman, S., & Naples, A. D. (2001). "Language is spatial": Experimental evidence for image schemas of concrete and abstract verbs. In *Proceedings of the 23th Annual Meeting of the Cognitive Science Society*. (pp. 873-878).
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: The MIT Press.
- Rogers, T. T., & McClelland, J. L. (2008). Précis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences*, *31*, 689-749.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental interpretation through contextual representation. *Cognition*, *71*, 109-147.
- Smith, L. B. (2009). From fragments to geometric shape changes in visual object recognition between 18 and 24 months. *Current Directions in Psychological Science*, *18*, 290-294.
- Smith, L. B. (2013) It's all connected: Pathways in visual object recognition and early noun learning. *American Psychologist*, *68*, 618-629.
- Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, *65*, 235-241.

- Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447-481.
- Staudte, M., & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human–robot interaction. *Cognition, 120*, 268–291.
- Sturt, P., Keller, F., & Dubey, A. (2010). Syntactic priming in comprehension: Parallelism effects with and without coordination. *Journal of Memory and Language, 62*, 333-351.
- Tabachnik, B., & Fidell, L. S. (2007). *Experimental design using ANOVA*. Belmont, CA: Duxbury.
- 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, 29*, 557-580.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science, 268*, 1632–1634.
- Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition, 108*, 51-68.
- Taylor, L. J., & Zwaan, R. A. (2008). Motor resonance and linguistic focus. *The Quarterly Journal of Experimental Psychology, 61*, 896-904.
- Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution, *Journal of Memory and Language, 39*, 558–592.

- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: studying online sentence processing in young children. *Cognition*, *73*, 89-134.
- Wassenburg, S. I., & Zwaan, R. A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *The Quarterly Journal of Experimental Psychology*, *63*, 1665-1670.
- Williams, L. E., & Bargh, J. A. (2008a). Keeping one's distance : The influence of spatial distance cues on affect and evaluation. *Psychological Science*, *19*, 302-308.
- Williams, L. E., & Bargh, J. A. (2008b). Experiencing physical warmth promotes interpersonal warmth. *Science*, *322*, 606-607.
- Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *The Psychology of Learning and Motivation*, *44*, 35-62.
- Zwaan, R. A. & Madden, C. J. (2005) Embodied sentence comprehension. In D. Pecher and R. A. Zwaan (Eds.) *Grounding Cognition. The role of perception and action in memory, language and thinking*. New York, NY: Cambridge University Press.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162-185.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology*, *135*, 1-11.
- Zwaan, R. A., Taylor, L. J., & de Boer, M. (2010). Motor resonance as a function of narrative time: Further tests of the linguistic focus hypothesis. *Brain & Language*, *112*, 143-149.

Curriculum Vitae

PERSONAL

Name Ernesto Eduardo Guerra Gil
Date of Birth 12th of December, 1982
Place of Birth Valdivia, Chile
Address Heerenstraße 32
45145, Essen
Email ernesto.guerra@uni-bielefeld.de

EDUCATION

1988-2000 Colegio Cahuala Insular / Liceo Armando Robles Rivera
Enseñanza Básica y Media

2002-2008 Universidad Mayor
Licenciatura en Psicología

2008-2010 Rijksuniversiteit Groningen
M.A. in General Linguistics
Universität Potsdam
M.Sc. in Clinical Linguistics

2010-2013 Universität Bielfeld
PhD Student in Linguistics