Hybrid models of cognition: The influence of modal and amodal cues in language processing tasks

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Eduard Berndt
aus Tiraspol/Moldawien

Tübingen
2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	30.10.2020
Stellvertretender Dekan:	Prof. Dr. József Fortágh
1. Berichterstatter/-in:	Prof. Dr. Barbara Kaup
2. Berichterstatter/-in:	Prof. Dr. Hartmut Leuthold
3. Berichterstatter/-in:	Dr. Tessa Warren

ABSTRACT

In recent years, a growing amount of evidence has been accumulated suggesting that at least some part of our cognition and especially language comprehension is embodied in actions, perceptions, and emotions, and is therefore multimodal in nature. While the debate in previous decades was focused on whether cognition indeed is embodied, today the discussion revolves more around the question when and how embodied representations are used and what their exact role is. The present dissertation is aimed at shedding light on this discussion, investigating the presence and the role of multimodal representations across different tasks and contexts. At first, a series of anagram-solving tasks investigating the influence of different modal cues on subsequent solving of anagrams of words associated with either the ocean (e.g., shark -> SARHK) or the sky (e.g., cloud -> CUOLD) was conducted. Combining a background picture depicting an ocean-sky scene with a shift of attention towards the upper half of the computer screen resulted in faster solution times for words associated with sky compared to words associated with ocean, while the reverse was true for a downward attentional shift. This finding was extended to emotional valence, using pictures either associated with a positive or negative emotional valence to prime words with a matching emotional valence. Indeed, anagrams were solved faster when the emotional valence of the picture matched the associated emotional valence of the solution word. Going back to the domain of vertical space, we tried to replicate the findings of the first set of experiments with another set of stimuli and the use of linguistic cues in the form of adjectives or sentences preceding the anagrams, paired with a vertical shift of attention. In contrast to pictorial cues, these linguistic cues did not influence solution times. In another set of anagram-solving experiments, we directly compared the influence of linguistic (amodal) and pictorial colour (modal) cues, using written colour words or coloured rectangles as primes for solution words associated with a certain colour (e.g., specific types of fruit or

vegetable, such as “cherry”). These were solved faster when a matching colour cue was presented before the anagram, regardless of whether the colour cue was linguistic or pictorial. Combining both cues by showing a written colour word inside a coloured rectangle only facilitated anagram solving of anagrams when both cues matched the solution word, e.g. the word “green” written inside a green rectangle facilitating solution of an anagram for “cucumber”. Neither a symbolic, amodal colour word, nor a colour patch seem to be responsible for this match effect exclusively, but instead both cues seem to activate the same superimposed conceptual colour representation. In a last line of research, it was investigated in how far hemispheric differences come into play during embodied word representations. A divided visual field study by Zwaan and Yaxley (2003a), who found a match effect regarding visual-spatial relations between objects to be confined to the right hemisphere in a semantic-relatedness judgment task was replicated - with the addition of the factor response side. Word pairs were shown very briefly either to the left or right visual field in a vertical arrangement on the screen either matching or mismatching the canonical spatial relation between the word’s referents (“nose” being above “mustache” in a canonical view of a face, thus seeing “nose” written above “mustache” would be a match). In contrast to the original study there was no interaction between visual field and the spatial compatibility effect. Instead an interaction between response side and visual field and an additional main effect of match - independent of visual field - was found. This leads us to assume that multimodal concepts are not confined to either hemisphere but instead seem to be spread over large scale networks across the whole brain. Taking all of these results together, a hybrid-view of cognition seems to be the most fertile: superimposed conceptual representations seem to be at the core of semantic meaning, and can be influenced by both modal and amodal contextual information, with neither type of information exerting clear dominance over the other.

TABLE OF CONTENTS

1 INTRODUCTION

Humans are the most intelligent and influential species on this planet. We create and use tools, with which we build cities, skyscrapers, cars and even airplanes. We can plan ahead and anticipate consequences of our own and others' actions, reminisce about the past and analyse previous events and evaluate them. We are able to influence and change our environment to our needs and (literally) reach for the stars. What is it that makes us as humans so special, that allows us to be able to achieve all of these and many more incredible feats? It can be argued that language is the single most important thing separating us humans from other animal species. While animals are also capable of different forms of communication, nothing comes close to the infinite amount of sensible utterances human languages possess, granting humans far superior communication skills. Language allows us to communicate our beliefs, ideas, memories or plans for the future. We are even able to do so without the contents of what we are communicating having to be in our physical vicinity for another person to understand what we mean, as long as we speak the same language. This allows us to in principle *experience* things that we would never be able to perceive ourselves in a single lifetime, or as Johnson-Laird (1983) put it, "language is experience by proxy" (*p*. 430). Thus, we are able to accumulate a vast amount of experience, our own and those of others, and pass it on to future generations, leading to an exponential growth of knowledge and ultimately to the wealth of technology we have today.

However, it is still not fully clear how language could evolve into what it is. Are we special because we possess the powerful tool that is our language, or did our ability to communicate and use language the way we do only evolve because we were special in the first place? How is meaning stored, processed and combined in our brain? Different theories tried to find different answers to these questions. Chomsky (1957, 1975) for example, assumed that humans possess some innate underlying structure, or proto-syntax, shared by all languages allowing us to produce sensible utterances and derive meaning from language. Thus, human cognition is based on innate cognitive structures and their creative use. Barsalou (1999, 2008) on the other hand assumes that cognition is rooted in interaction with our environment, combining experiential features into concepts. These concepts are also tightly linked to language and thus can activate words corresponding to the concept in question, or certain words can activate their corresponding concepts and the experiential features linked to it.

While there is evidence for both accounts, the question how meaning is ultimately stored and used in cognitive processes remains unresolved. In the following chapters, after outlining different theories of human cognition, a series of experiments is reported and the results as well as their implications for the current state of research on meaning representation are discussed.

1.1. Traditional and connectionist theories of cognition

Traditional theories of cognition emerged at the beginning of the second half of the last century along with the cognitive revolution in psychology and propose that language is processed and stored in a more or less isolated language module (Chomsky, 1957, 1975, 1980; Fodor 1975, 1983; Pylyshyn, 1989; Fodor & Pylyshyn 1988). The so-called *language faculty* is described by Chomsky (1957, 1975, 1980) as being an innate, genetically determined mental structure, akin to a mental organ in our brain and is supposed to enable us to learn and understand human language. This mental organ *grows* under normal conditions to allow us to rapidly learn a language with its complex systems of rules (distinction between verbs and nouns, distinction between function words and content words and reciprocal relations) which would not be possible without such a specialized module. Since human languages operate entirely on abstract symbols (e.g. letters and words), traditionalists further argued that language supposedly mirrors our mental conceptual representations and assumed that these knowledge representations must work in a similar way. This led to what Fodor (1975) called *language of thought*. Traditional theories assumed that this mental language needs to be based on abstract symbols just as spoken languages, to allow for truth-preserving logical operations like productive combinations, the representation of propositional relations, type-token relations and the production of inferences (Fodor & Pylyshyn 1988). Upon perceiving an entity or an event in the world, perceptual states are supposed to be transduced into these abstract symbols that make up conceptual representations. This means that conceptual representations themselves are arbitrary, similar to how most words do not resemble their referent in any meaningful way, and amodal, since they are non-perceptual in nature (Barsalou, 1999).

Starting in the early 1980s, traditional theories were challenged by connectionist models (e.g. McClelland & Rumelhart, 1986). Although connectionists did not directly challenge the assumption that knowledge structures are abstract and amodal, they claimed that no innate mental structures are necessary to explain the complexity of language and knowledge representations. Instead, concepts or propositions represent nodes in a network that are interrelated. Activation of a node spreads along the connections of the network and can in turn activate or inhibit other nodes. Connections between these nodes are strengthened or weakened based on simple rules and interactions with the environment to form complex knowledge structures in a bottom-up manner.

One influential example of a computationally implemented model of how humans create symbolic meaning representations is the Construction-Integration Model from Kintsch (1988), which evolved from earlier versions of similar models (Kintsch & van Dijk, 1978; Kintsch, 1985). This model mainly focuses on discourse comprehension, namely on how the meaning of a text is transformed into a conceptual representation. Two vital steps are proposed: First, a knowledge base is constructed using sloppy inference rules resulting in an associative network. This initial knowledge base may well be incoherent, since many different associations and inferences have been activated. Only in the second step, the integration phase, a coherent conceptual representation can be formed by strengthening fitting or weakening unfitting associations, depending on the context. The nodes of these networks take on the form of propositions that describe relations between concepts or concepts and other propositions. They consist of a head and a number of slots that can be filled with different arguments. To make it clearer, a sentence like "John loves Mary." could be represented as a propositional representation in the form of LOVE [JOHN, MARY], defining JOHN as the agent and MARY as the object of the relation LOVE[1]. These propositions can become more and more complex to represent more complex sentences, for example: "John is happy, because Mary loves him." could be transduced into a propositional representation of the form HAPPY [JOHN, BECAUSE [LOVE [MARY, JOHN]]]. Here, the main propositions head is HAPPY, the agent slot is filled with JOHN. Another proposition follows with the head BECAUSE to specify the reason for John's happiness, which turns out to be the proposition LOVE with MARY in the agent role and JOHN as the object of the relation. While reading the text, more and more propositional representations are constructed and integrated into a text base of the meaning of the whole text. Missing information can be inferred from other knowledge propositions that have not been explicitly mentioned in the text but are closely associated with the propositions that are part of the current text base. In several processing cycles this knowledge base is modified and expanded upon reading new sentences. Propositions that are not referenced in later parts of the text do not contribute to the coherence of the knowledge base and are classified as irrelevant and deleted. Ultimately a complete and coherent propositional knowledge structure is built that represents the gist of the text that was read in the form of a network of propositions.

While traditional theories of language representation and the early computational models seem to correctly capture different parts of language processing, like the ability to represent types and tokens, produce categorical inferences, represent propositions and productively combine symbols (see Barsalou, 1999), it is not really clear how meaning is ultimately derived from these knowledge structures. Although it is not mentioned explicitly,

---

[1] Note that from now on, words written in capital letters will indicate a concept corresponding to the word, instead of the word itself.

the arguments used in Kintsch (1988) are likely assumed to take on the form of abstract lexical representations, as were introduced by traditional theories of language representations. The important question though is how and why representations are transduced from perceptual states to amodal symbols and then possibly back again. How is meaning constructed when only using abstract, amodal and arbitrary symbols? What good is a propositional knowledge structure consisting of arguments when these arguments are basically hollow shells?

To make this argument clearer, Searle (1980) made use of a thought-experiment: Imagine a non-Chinese-speaking human sitting inside a closed room, containing Chinese scripts and a manual of certain rules written in his native language. Another Chinese person gives him a Chinese text and questions about this text. Using the manual and the Chinese scripts, the human could theoretically answer the questions and write them on another sheet of paper by comparing the symbols and using the rules he has at his disposal. The Chinese person would think that the person inside the room did actually understand Chinese to answer the question, however all he did was use the rules he was given without knowing anything about the meaning of the Chinese symbols. Applied to mental representations, this would mean that we are able to form associative networks that resemble knowledge structures consisting of interconnected symbols, but ultimately never know anything about the real world that the symbols in the network are supposed to represent. This has become known as the grounding problem, described by Harnad (1990). How is meaning derived from symbols if they are only related to each other and completely separate from our perception of the real world? A possible solution to this problem was proposed by embodied theories of cognition described in the next section.

1.2. Embodied theories of cognition

Since the end of the last century, embodied theories of cognition have emerged, challenging many of the proposed claims of traditional theories presented in the previous section (e.g. Barsalou, 1999, 2008; Glenberg & Kaschak, 2002; Zwaan & Madden, 2005). Instead of assuming that a perceptual state resulting from the perception of an entity or an event in the real world is transduced into another amodal representation, Barsalou (1999) argued that these perceptual states themselves could actually be what constitutes mental conceptual representations. This means that upon hearing or reading a word like "chair", the same neural assemblies that were active during the perception of chairs are reactivated to capture the meaning of "chair". This *neural reuse* theoretically solves the grounding problem (Harnad, 1990) as described in the previous subsection, since mental conceptual representations would be grounded in the actual perceptual states corresponding to entities or events in the real world. Such perceptual states can be *simulated* mentally, allowing the re-enactment of perceptual states even in the absence of their real world counterparts. This effectively constitutes such perceptual symbol system as a conceptual system instead of reducing it to merely a recording system.

Barsalou (1999) went on to argue that perceptual symbols represent small components of different perceptual aspects of entities (like shape or orientation) or events (like temporal order or cause and effect) and further include introspective states (like emotions). This allows perceptual symbols to satisfy the most important aspects of a conceptual system according to traditional theories: Perceptual symbols can be combined productively, meaning that the conceptual representation for "red chair" for example could be evoked by the activation of perceptual properties of RED and combining them with the conceptual representation of CHAIR, resulting in a chair that has the colour red and contains both the perceptual symbols for RED and CHAIR. Similarly, propositions could be built using only using perceptual symbols by establishing type-token relations between certain concepts and real world agents. Imagine the concept HUNGRY which could consist of different perceptual symbols representing its meaning, mainly introspective states when oneself was hungry. This concept could be simulated and applied to an agent like John, who is sitting next to you and just told you that he is hungry. Projecting your own concept of HUNGRY onto John, allows you to effectively represent the proposition HUNGRY(JOHN). Furthermore, it enables you to draw inferences based on the simulation of the concept in question, for example that John probably wants to go and get something to eat, since this is probably what you associate with being hungry. So instead of assuming the existence of amodal symbols that are transduced from perception, for which there is basically no evidence, a conceptual system may actually be fully grounded in perception, action, emotion and personal experiences.

Zwaan and Madden (2005) built on these ideas and described what they referred to as an experiential simulation during language comprehension. They assume that upon reading or hearing a text readers or listeners engage in a mental simulation of the linguistic content. This simulation is achieved by reactivating multimodal experiential traces that are associated with the linguistic input through co-occurrence. Because these simulations should be multimodal in nature, Zwaan and Madden (2005) claim that they should contain perceptual aspects of referents or situations, spatial relations between objects or object parts, dynamic aspects of events as well as perspective. Indeed, there have been empirical studies supporting these ideas. Zwaan, Stanfield and Yaxley (2002) for example tested the assumption that perceptual aspects of referents are activated upon reading a sentence. Participants read sentences like “John saw an egg in the fridge/frying pan” and subsequently saw a picture of a whole egg as it would look in a box in the fridge, or a picture of a fried egg. When the picture and sentence matched (fridge and picture of a whole egg or frying pan and picture of a fried egg), participants were faster to correctly state whether the object shown in the picture was mentioned in the sentence compared to trials with mismatching sentence and picture.

Looking at representations stemming from single words instead of sentences, Zwaan and Yaxley (2003b) were able to find evidence for the activation of spatial relations between objects: Showing a word pair with one word appearing above the other in a semantic-relatedness judgement task, they found a match effect with faster response latencies when the arrangement of the words on the screen mirrored their canonical spatial relation to each other. When “nose” is written above “mustache” for example, it would be in accordance with an assumed experiential simulation upon reading both words, while “mustache” written above “nose” would be in conflict with it. Thus, the authors saw this result as evidence for the existence and involvement of multimodal experiential simulations in language processing. Zwaan, Madden, Yaxley and Aveyard (2004) found evidence for the remaining two assumptions of Zwaan and Madden (2005) regarding experiential simulations, namely simulation of dynamic aspects of events as well as simulation of perspective. Participants heard sentences and judged whether two sequentially presented visual objects were the same. Critical trials consisted of sentences implying a movement towards or away from the observer, for example “The pitcher hurled the baseball towards you” or “You hurled the baseball towards the batter”. Subsequently, two pictures of the critical object of the sentence were shown, a baseball in this case, with the second picture being slightly larger or smaller than the first picture, reflecting movement of the ball towards or away from the reader. Participants were faster to respond when the change of object size on the pictures was congruent with the situation described in the preceding sentence compared to incongruent trials.

While the studies discussed so far have focussed on perceptual information, there have also been numerous studies showing involvement of motor areas during language processing. Glenberg and Kaschak (2002) found match effects between sentences describing a movement away or toward one’s body like “He closed/opened the drawer” and a required motor response involving movement towards or away from the body of the participants in a sensibility judgement task. These results were interpreted to be consistent with the idea that meaning is based on action, action-based goals and affordances of objects in our environment (see also Glenberg & Gallese, 2012).

Indeed, there have been several other studies showing involvement of motor areas during language processing. Hauk, Johnsrude and Pulvermüller (2004) found that areas in the motor and premotor cortex were active during processing of action verbs like “lick”, “pick” or “kick”. The activation found included activation of the motor strip either adjacent or overlapping the brain areas that are actually involved during movement of the face, hands or feet, depending on which of the verbs was being processed. Lachmair, Dudschig, de Filippis, de la Vega and Kaup (2011) found match effects between the associated vertical position of a word’s referent (“airplane” is associated with an upper vertical position, while “cellar” is associated with a lower vertical position) and subsequent upward or downward movement of the one’s arm; the same has been found for verbs indicating an upward or downward movement (Dudschig, Lachmair, de la Vega & Kaup, 2012).

However, one of the biggest criticisms of embodied accounts of cognition has always been the question how abstract concepts with no actual perceptual analogue (for example “love” or “justice”) can be represented on the basis of perceptual symbols. While it is currently still a topic under discussion, Barsalou (1999) already suggested that “Abstract dimensions are grounded in complex simulations of combined physical and introspective events.” (p. 577). In the following years there have been manifold suggestions regarding the grounding of abstract words, for example via the metaphoric mapping account (Boroditsky, 2000; Lakoff & Johnson, 1980), the grounding of abstract words in emotions (Kousta, Vigliocco, Vinson, Andrews & Del Campo, 2011), the grounding of valence in space (e.g., Meier & Robinson, 2004) or body-specific associations between fluency and valence (e.g., Cassasanto, 2009). In line with this assumption, match effects between object properties that are generally deemed abstract and motor responses have indeed been found. For example, de la Vega, De Filippis, Lachmair, Duschig and Kaup (2012) found a match effect between emotional valence of presented words and response hand. While right-handers were faster to respond to words of positive valence with the right hand, it was the opposite for left-handers, indicating a mapping of positive valence to one's dominant hand. Further examples include the abstract concept of time that seems to be linked to the concept of space, with the future being linked with either front or right and the past with either back or left (Ulrich et al. 2012; Eikmeier, Alex-Ruf, Maienborn & Ulrich, 2015; Santiago, Lupáñez, Pérez & Funes, 2007). Additionally, it was found that implied pitch height during sentence comprehension was also found to be linked to space (Wolter, Dudschig, de la Vega & Kaup, 2015).

To conclude, there have been numerous studies providing evidence for the involvement of sensorimotor systems during language comprehension, arguing against the claim of traditional theories of cognition that conceptual knowledge is based on abstract and amodal representations and that language is processed in a specialized language module in the brain that is more or less independent of other subsystems.

1.3. Hybrid theories of cognition

As described in the previous section, there has been an impressive amount of empirical evidence for the involvement of perceptual systems during language comprehension and other cognitive tasks supporting the claims of embodied cognition accounts. This has led to the widespread acceptance of the assumption that our sensorimotor systems are involved in language processing. However, there are also several core questions that cannot be answered by the embodied accounts as outlined today. For example, Pecher, Boot and Van Dantzig (2011) pointed out that a basic metaphoric mapping account – mapping for example positive to the upper space and negative to the lower space – ignores way too much of individual meaning aspects of the represented concepts, that it could account for a full meaning comprehension process. In other words, how can the words "love" and "happiness" being discriminated, if all information we have to represent them is their association with the upper space? Beyond this issue of how the present evidence in favour of embodied meaning representations can account for the diverse meaning aspects that are there to represent, there is also an ongoing debate about the actual role and importance of these sensorimotor activations (Mahon & Caramazza 2008, Goldinger, Papesh, Barnhart, Hansen & Hout, 2016). Non-embodied accounts of language comprehension have adapted to include sensorimotor systems, like the "faculty of language in a broader sense" described by Hauser, Chomsky and Fitch (2002). Many embodied accounts acknowledge the need for some kind of amodal system or representation, resulting from gradual abstraction of information from perceptual and motor systems (Schuil, Smits & Zwaan, 2013; Zwaan 2014, Dove 2009). These so-called hybrid accounts of cognition assume the existence of both multimodal and abstract conceptual representations, but can differ greatly in regard to the importance that is ascribed to either system. Basically, they can be sorted on a continuum from a very weak form of embodiment (Mahon & Caramazza, 2008, Patterson, Nestor & Rogers, 2007) to strong embodiment (Glenberg & Kaschak, 2003; Gallese & Lakoff, 2005). Extensive reviews on this matter were provided by Binder and Desai (2011), Chatterjee (2010) and Meteyard, Cuadrado, Bahrami and Vigliocco (2012), among others.

Most of these hybrid accounts agree on the existence of two more or less separate systems, a sensorimotor system as well as an abstract symbol system, something that has already been proposed by Paivio (1986). Importantly however, they vary in their assumptions about the respective importance of these different systems to knowledge representation and what actually constitutes core concepts. Chatterjee (2010) emphasizes the importance to not jump to conclusions in favour of embodied accounts of cognition too quickly, and states that "A focus on disembodying cognition, or on graded grounding, opens the way to think about how humans abstract." (Chatterjee, 2010, p. 79). He calls for more research on how exactly we are able to abstract perceptual or motor input into linguistic units and vice versa, and how and where these processes take place in our brain, specifically mentioning potential laterality differences in content and structure of association cortices between the left and right hemisphere. Binder and Desai (2011) as well as Meteyard et al. (2012) conclude from a review of empirical and neuroimaging data that our core concepts are likely stored or processed in so-called *convergence zones* located at convergences of different perceptual processing streams, in line with Damasio (1989). These zones are not modality-specific but instead integrate and process information from different modalities ultimately forming what Binder and Desai (2011) call a *supramodal* representation. While these representations allow for different levels of abstractions, their core input is information passed on from perceptual and motor systems.

Mahon & Caramazza (2008) on the other hand assume that at their core, concepts are abstract and symbolic but can employ perceptual and motor systems to gain more information on the look, feel or use of their referents for example. Since concepts are not constituted by multimodal representations, the empirical evidence is interpreted in the light of spreading activation of these abstract conceptual cores to adjacent or associated perceptual and motor systems and a possible interference between the task or the context and the retrieval of relevant information by the core concept. They argue for a disembodied cognition, harshly criticizing embodied accounts of cognition (also see Leshinskaya & Caramazza, 2016; Mahon, 2015), although most of the criticism applies only to particularly strong versions of embodiment (e.g. Gallese & Lakoff, 2005; Glenberg & Kaschak, 2002), for example the claim that concepts are entirely reducible to modality-specific sensory or motor representations.

To conclude, there is mostly agreement on the fact that perceptual and motor systems are in some way involved in cognition and language processing as are more amodal representations and processes. This would speak in favour of hybrid models of cognition. In the following I will discuss in some more detail two lines of empirical research, that are typically interpreted as supporting this hybrid view, namely research on context dependency and on the functional role of sensorimotor processes for comprehension.

1.4. Current state of the debate and open questions: Context Dependency and Functional Role

As we have seen in earlier chapters there are three classes of representational models heavily discussed in the language domain. First, propositional models basing language processes solely on abstract, amodal symbols. Second, embodied models claiming the core processes of language are based on sensorimotor, modal representations. And finally, hybrid models postulating two types of representational formats. There is a large amount of empirical evidence in favour of the involvement of sensorimotor processes to language processing. Nevertheless, there are core open debates with regard to representational issues in the language domain, namely concerning the context dependency of the effects attributed to the involvement of sensorimotor representations and the interrelated issue regarding the functional relevance of these types of representations for comprehension. First, I will turn to the debate regarding the context dependency, second the issues regarding functional relevance will be introduced.

1.4.1. Context Dependency

Interestingly, many results in the embodiment literature pointed towards a strong context dependency of the activation of modal representations, which in principle fits well with the hybrid account. Context dependency has been observed with respect to task-requirements such as processing depth, aspects of the stimuli- and response-sets as well as hemispheric specificities regarding the activation of modal representation during language representations. These findings will be outlined in detail below.

First, with regard to task dependencies: Indeed, some of the effects described earlier are only found for specific tasks, like the mapping of time to space (see Ulrich et al., 2012), which was present when participants responded to whether the sentence at hand was past- or future-related but not when the task was a sensibility judgment. Similarly, the mapping of positive valence to dominant hand was only apparent when participants were told to explicitly answer to the valence of words (see de la Vega et al., 2012). Also when investigating gestural knowledge evoked by words, it has been shown that a task that demands attention to the word meaning is required in order to evoke sensorimotor representations to be activated (Bub, Masson & Cree, 2008).

Second, with regard to properties of the stimuli and/ or response sets: The influence of associated vertical position on subsequent motor responses has been shown to require that the vertical dimension is salient in either the stimulus or the response set, indicating that this association is not activated automatically (Dudschig & Kaup, 2017; Areshenkoff, Bub & Masson, 2017). Interestingly, Gozli, Chaasten and Pratt (2013) showed that the inclusion of different stimulus sets and word categories sometimes even results in reversed effects.

A final specificity of the involvement of rather modal representations can be found in hemispheric differences reported regarding the activation of conceptual representations. There is evidence that points to hemispheric differences regarding conceptual processing, possibly reflecting the distinction between a multimodal and an amodal conceptual system (Paivio & Ernest, 1971). It has been shown for example, that the right hemisphere is involved in the processing of visual-spatial relations (Marsolek, Kosslyn & Squire, 1992), while the left hemisphere seems to responsible for more abstract associations and relations like category membership (Abernathy & Coney, 1996). A particularly clear result regarding hemispheric differences and the involvement of modal representations was provided by Zwaan and Yaxley (2003a). In their study the authors used a divided visual field technique and reported that the influence of modal spatial relationships on language processing is confined to the right hemisphere. If stimuli are only presented in a manner that results in initial left hemispheric language processing, no such modal influences on language representation are found. On the other hand, there are also studies clearly showing that embodied language representations evolve in large semantic network spanning the whole brain (Binder & Desai, 2011). This network can be divided into different sections mostly reflecting different modalities like vision, sound or emotion, while more abstract representations seem to be located in convergence zones between these modal systems. However, no clear distinction between the hemispheres regarding modal and amodal representations seems to exist.

Taken together there is ample evidence in the literature that the involvement of sensorimotor representations during language comprehension is strongly language dependent. As indicated above this evidence alone could be seen as support for the hybrid models of comprehension. However, in principle this evidence would be in line with a model where sensorimotor processes are not functionally relevant for comprehension, but rather are an optional by-product of comprehension. According to this view comprehension processes would be based on amodal abstract representations (as assumed by traditional theories and outlined in Mahon & Caramazza, 2008; Mahon, 2015), whereby sensorimotor representations would be activated if the capacities are available or the task or stimulus and response sets call for such activations in a resonance-like manner. A true hybrid model would need to assume that the sensorimotor representations do play a functional role for comprehension under certain conditions. In the following paragraph I will therefore report research directly addressing this issue.

1.4.2. Functional Relevance

A large amount of studies have provided evidence for the involvement of sensorimotor representations during language processing. The previous paragraph indicated that some of these findings are context dependent. This leads to a way bigger question: What is actually the role of these types of modal representations? Are they functional for comprehension? This question is particularly relevant for the issue of what type of language processing model is actually valid. As outlined in the previous sections different models of language comprehension vary with regard to their claim regarding the involvement of modal representations in the comprehension process. If indeed hybrid models are correct in their assumption as to how language is processed, than one would assume that both modal and amodal representations do play a functional role for comprehension.

The research into the functional role of sensorimotor representations for the language comprehension system has been rather neglected in the literature so far. Indeed, when looking at the literature showing the involvement of sensorimotor representations during language comprehension there is one astonishing fact: With regard to direction of influence it has been predominantly shown that linguistic processing influences the subsequent processing of sensorimotor information, however, there seems a lack of evidence showing that modal cues influence subsequent language processes. This type of evidence is particularly important, as it actually looks into the comprehension processes and thereby excludes the option that the activation of sensorimotor processes are merely a by-product of comprehension processes.

The few studies addressing this issue came to diverse conclusions. Studies by Kaschak and collegues reported evidence for effects of visual and auditory motion (towards, away, up and down) on sensibility judgment times of sentences (e.g., “The cat climbed the tree”) (Kaschak et al., 2005; Kaschak, Zwaan, Aveyard & Yaxley, 2006). However, whether a match or mismatch advantage was observed strongly depended on the modalities used. Also Meteyard, Zokaei, Bahrami, and Vigliocco (2008) reported mixed evidence regarding the influence of visual motion on lexical decision times: only when visual motion (dot patterns moving up or down) was presented at near-threshold levels an influence on lexical decision times of verbs (e.g., “rise” and “fall”) was observed. Beyond the influence of perception on action some studies also looked at the effects of action on language processing. For instance, Glenberg, Sato and Cattaneo (2008) found that repeatedly performing a directional arm movement (e.g., moving beans away or towards one’s body) before a comprehension task resulted in slower sensibility judgment times if the arm movement matched the movement implied by the sentences. It should be noted that this result is not easy to interpret considering that a repeated arm movement away from one’s body involved the same amount of arm movements towards one’s body. In contrast to this finding, Pulvermüller, Hauk, Nikulin and Ilmoniemi (2005) used TMS techniques to stimulate hand or leg sites in the brain, which resulted in faster responses to matching verbs (e.g., “pick” or “kick”). Also Rüschemeyer, Lindemann, van Rooij, van Dam and Bekkering (2010) investigated the effects of a motor task on lexical decision times. Words referring to objects that demand an action affordance (e.g., “cup”) were responded to more accurately than words not evoking such affordances when accompanied by a motor task. Finally, Strozyk, Leuthold, Miller and Kaup (2018) across five experiments found no influence of action preparation on lexical decision times when investigating spatial terms (e.g., “sun” vs “worm”), even though it could be shown that participants indeed prepared the responses.

In sum, these findings are rather sparsely to find, which is a clear opposition to the amount of studies reporting the corresponding influence in the other direction, namely from language processing on subsequent sensorimotor processes. This is specifically surprising as these types of findings can be regarded as a pre-condition for sensorimotor representations being functionally relevant for the comprehension processes. Also relevant to this question are studies looking into whether comprehension is hampered when comprehenders are prevented from simulating, for example by occupying relevant sensorimotor systems. For instance, a study by Strozyk, Dudschig and Kaup (2017) investigated whether a secondary tapping task involving either the hand or the feet – and thereby occupying the respective systems - specifically impairs comprehension of hand-related (e.g., “cup”) or foot-related words (e.g., “shoe”) in a lexical decision task. They found no effector specific impairment by the secondary task and concluded that in this setting effector-specific activations were not functional for solving this type of comprehension task. In contrast Yee, Chrysikou, Hoffmann and Thompson-Schill, 2013 (2013) found a specific influence of a secondary tapping task engaging the hands on a classification of nouns referring to object that participants have experience manipulating. No such influence was found for a non-manual mental rotation task. Thus, also this area of investigation is at the moment only addressed in very few studies and the conclusions are rather mixed. Again this fits best with hybrid models of comprehension postulating two representational formats that are functionally relevant under different conditions.

1.5. The aim of this thesis

The discussions in the previous sections regarding the functionality and context dependency seem to clearly suggest that accounts postulating only one representational format – be it either abstract amodal or sensorimotor modal – are misled. In contrast, the results are well in line with hybrid models. In the present thesis the hybrid view of language comprehension shall be further investigated. Given the relevance of the functionality of modal representations for comprehension – as a real hybrid model only distinguishes itself from models that assume modal representations are an optional by-product by assuming they are indeed functionally relevant - this question shall be further investigated in the current thesis. This will be done by means of the influence of modal cues on anagram solving times. If indeed both representations do play a role it would be important to investigate whether one type of representation is predominant – this will be done by investigating whether modal cues have an advantage over amodal cues or the other way round in an anagram solving task. And finally, this thesis aims at contributing to the question under which conditions one or the other representational format might be predominant. Therefore the previously introduced suggestion regarding hemispheric differences and representation format will be further investigated.

Taken together this thesis will address three questions central to any hybrid account of comprehension. (1) Are modal representations functional for comprehension in certain domains. (2) Under which conditions do modal or amodal representations play a predominant role.(3) Are hemispheric representation specificities one key to differences in representational format as postulated by Zwaan and Yaxely (2003a).

2. EXPERIMENTS

2.1. Influence of modal and amodal cues on anagram solving times

To investigate the influence of multimodal experiential traces on the activation of conceptual representations, we decided to employ anagram solving tasks (Berndt, Dudschig & Kaup, 2018a, 2018b). An anagram is a series of scrambled letters that form a word or sentence when put in the right order and anagram solving tasks are mostly found in problem solving literature. Anagrams can either be solved by rearranging the letters until a sensible solution is found, this approach is called search solution, or the correct answer can come to mind suddenly, called pop-out solution. The latter relies heavily on gradual accumulation of partial information, with usually the letters serving as cues, until a lexical representation becomes active (Ellis, Glaholt & Reingold, 2011). When additional cues are provided, for example associated words serving as primes ("table" shown before an anagram for "chair"), anagram solving is facilitated (Dominowski & Ekstrand, 1967). Following this logic, we decided to investigate the influence of different modal cues on solving times of subsequently presented anagrams, actually measuring the pre-activation of a given conceptual representation. When the solution word and the modal cue share perceptual features, these cues should activate parts of the conceptual representation of the solution word. While this conceptual representation is unconscious at first, the modal cues contribute to the gradual accumulation of information associated with this conceptual representation, along with the letters of the anagram, until it becomes conscious. This should lead to faster access of the underlying lexical representation and in turn lead to faster anagram solving times.

2.1.1. Spatial meaning domain: Attention shifts and modal background pictures as cues

As mentioned in the introduction, a word's associated vertical position can facilitate subsequent upward or downward arm movements (Lachmair et al. 2011; Dudschig et al. 2012). Šetić and Domijan (2007) additionally found that lexical access of words describing flying animals was facilitated when they were presented at the top of the screen, while words describing non-flying animals were recognized faster when shown at the bottom of the screen. Similarly, Pecher, Van Dantzig, Boot, Zanolie and Huber (2010) presented words associated with the sky or the ocean at the top or at the bottom of the screen and participants had to judge whether what was described by the word could be found in the sky or in the ocean. Participants were faster to decide whether something belonged in the sky when the word was presented at the top of the screen and faster to decide whether it belonged in the ocean when it was shown at the bottom of the screen.

Building on these findings we conducted a series of experiments, in which we we used 60 nouns that were either associated with the ocean (for example “dolphin” and “shipwreck”) or the sky (for example “airplane” or “eagle”) and used them in an anagram solving task after scrambling the letters (Berndt, Dudschig & Kaup, 2018a). Furthermore, we used positional information, pictorial information or both, serving as modal cues, to investigate their influence on the availability of the solution words. The hypothesis was that participants should be faster to solve the anagrams when congruent modal cues were shown before the anagram, similarly to associative priming with words associated with the solution words (Dominowski & Ekstrand, 1967). If conceptual representations are indeed at least in part multimodal, these modal cues should have the same effect as associated words insofar as they should pre-activate the concept just the same. In turn, lexical access to the solution word should be facilitated.

In the first experiment the anagrams appeared either at the top or at the bottom of the screen after a fixation cross was shown in the centre of the screen. The resulting attentional shift to the upper or lower vertical edge of the screen served as a vertical cue (see Šetić & Domijan, 2007 and Pecher et al., 2011). To test our hypothesis, we used linear mixed effect models to find out whether the factor congruency would significantly improve the model fit. Congruent trials, with an anagram of a sky-word appearing at the top of the screen or an anagram of an ocean-word appearing at the bottom of the screen, did not significantly facilitate anagram solving times compared to incongruent trials, using vertical cues alone neither for reaction times, $\chi 2(1) = 0.003$, $p = .96$, $\beta = 63.84$, $t = 0.07$, nor for the accuracy rates, $\chi 2(1) = 0.45$, $p = .50$, $\beta = 0.013$, $t = 0.66$. In an attempt to further emphasize the vertical dimension, we decided to include a pictogram of a human in the middle of the screen for the second experiment. Anagrams appeared either above or below this pictogram of a human. Congruent trials were still not solved faster or more accurate than incongruent trials (reaction times, $\chi 2(1) = 0.02$, $p = .87$, $\beta = -118.2$, $t = -0.15$; accuracy rates, $\chi 2(1) = 0.14$, $p = .70$, $\beta = -0.006$, $t = -0.38$), showing that vertical cues alone did not sufficiently pre-activate conceptual representations of the solution words. In fact, the pictogram could have hindered the activation of multimodal representations, because it was not helpful in engaging in an experiential simulation of the solution words, since shipwrecks are usually not experienced below a human standing upright. To account for this, we decided to include background pictures that emphasized the vertical dimension and further matched the ocean and sky theme of the solution words in the third and fourth experiment. A line drawing of a sail boat was shown in the centre of the screen in the third experiment and a significant influence of the congruency factor was found, with congruent trials being solved faster than incongruent ones ($\chi 2(1) = 4.19$, $p < .05$, $\beta = 1677$, $t = 2.07$). However, it did not significantly influence the accuracy rates ($\chi 2(1) = 0.007$, $p = .94$, $\beta = -0.001$, $t = -0.08$). To further solidify and replicate this finding, we used a background picture of a horizon, showing the ocean at the bottom half of the screen and the sky at the top half of the screen. The same pattern of results emerged when we used this more realistic background picture that filled the whole screen, with faster solving times for congruent trials compared to incongruent ones but still no influence on accuracy rates (reaction times: $\chi2(1) = 4.40$, $p < .05$, $\beta = 2519.6$, $t = 2.12$; accuracy rates, $\chi2(1) = 0.15$, $p = .70$, $\beta = 0.008$, $t = 0.38$). An overview of the methods and solving times is depicted in Figure 1.

To rule out two possible alternative explanations to this pattern of results that would be more in line with the assumption that meaning representations are abstract and amodal, we conducted two additional control experiments. It could have been possible that we only found significant congruency effects when showing a background picture that matched the ocean and sky theme, because it could have led participants to pre-activate abstract category concepts for sky and ocean, resulting in faster access to words belonging to one of these categories. Showing the anagrams at the top or at the bottom of the screen could have activated abstract concepts for TOP and BOTTOM respectively, further facilitating access to possible solution words. To test this, we replicated our first experiment, with the addition that participants were told before the experiment, that all possible solution words will either be part of the sky or ocean category. Although solution times were higher overall compared to the other experiments, congruency did not have a significant influence (reaction times: $\chi2(1) = 0.03$, $p = .86$, $\beta = 115.1$, $t = 0.19$; accuracy rates: $\chi2(1) = 1.14$, $p = .29$, $\beta = 0.021$, $t = 1.06$). Another possible explanation is that participants focused mainly on the part of the screen where the anagram appeared, which in turn semantically primed words associated with what is seen in that part of the screen. Looking at the sky in our fourth experiment for example, could have activated words associated with SKY and looking at the ocean in turn could have activated words associated with OCEAN. In this case, there is no need to assume that participants engaged in a perceptual simulation because these concepts could have been completely amodal. To rule out this second alternative explanation, we divided the background picture used in our fourth experiment, creating one

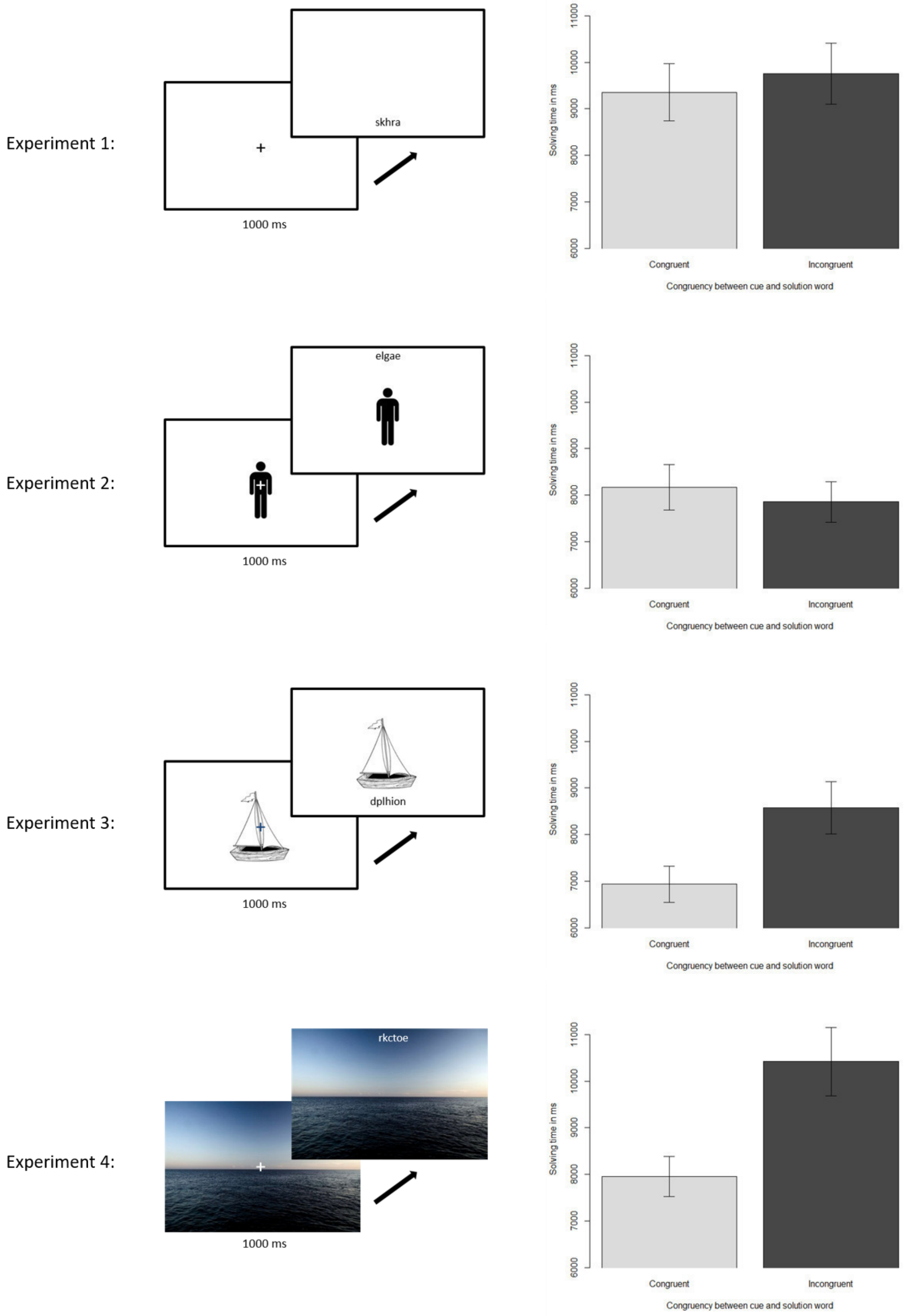


Figure 1. Sketch of the procedure and the results of the first four experiments in Berndt, Dudschig and Kaup (2018a). Error bars in this and following figures indicate 95% confidence intervals according to Masson and Loftus (2003).

Figure 2. Procedure and results of the sixth experiment in Berndt, Dudschig and Kaup (2018a).

background picture showing only the sky and another background picture showing only the ocean. Anagrams were presented centrally on either of these background pictures. Again, no significant influence of congruency was found as can be seen in Figure 2 (reaction times: $\chi 2(1) = 0.18$, $p = .68$, $\beta = -255.4$, $t = -0.42$; accuracy rates: $\chi 2(1) = 0.04$, $p = .85$, $\beta = -0.004$, $t = -0.19$).

This experiment addressed research question one of this thesis concerning the issue of functional relevance by means of investigating the reverse influence of modal activations on comprehension processes. The results are in line with our hypotheses and with the idea that multimodal experiential traces can activate meaning representations and in turn facilitate lexical access. This is an important addition to the previous literature, showing that not only do words activate modal traces like perceptual aspects associated with these words (e.g., Zwaan, Stanfield & Yaxley, 2002) or motor activation (e.g., Lachmair et al. 2011), but a sufficient amount of modal traces can also activate words associated with them.

2.1.2. Valence meaning domain: Modal background pictures as cues

In an attempt to generalize the congruency effect found in Berndt, Dudschig and Kaup (2018a) to a different set of stimuli, we employed a similar procedure to words that were associated with either a positive or negative emotional valence (e.g. “Love” and “War” respectively). As mentioned briefly in the introduction, de la Vega, De Filippis, Lachmair, Dudschig and Kaup (2012) found an influence of the emotional valence of a word and response hand, with positive words eliciting faster right-handed responses and negative words leading to faster left-handed responses in right handers. Following the same logic as in the previous experiments, we wanted to test whether we could facilitate the lexical access of positive or negative words by inducing positive or negative emotions. We used pictures from the IAPS (Lang, Bradley & Cuthbert, 2008) associated with either positive or negative emotions (e.g. a picture of a happy family eating dinner together or a picture of a sick and wounded animal respectively) and anagrams of words rated regarding their emotional valence in an online study to test this. The content of the pictures did not have semantic similarities to the subsequently presented solution word of the anagram except for the emotional valence to rule out that the pictures directly primed the solution word. The results indeed showed a significant congruency effect (χ^2 (1) = 12.44, $p < .001$), with congruent trials being solved faster compared to incongruent ones (see Figure 3). This finding further supports the idea that meaning representations can be activated by perceptual cues and subsequently facilitate lexical access to words that share a property with the pre-activated conceptual representation, namely emotional valence as tested here.

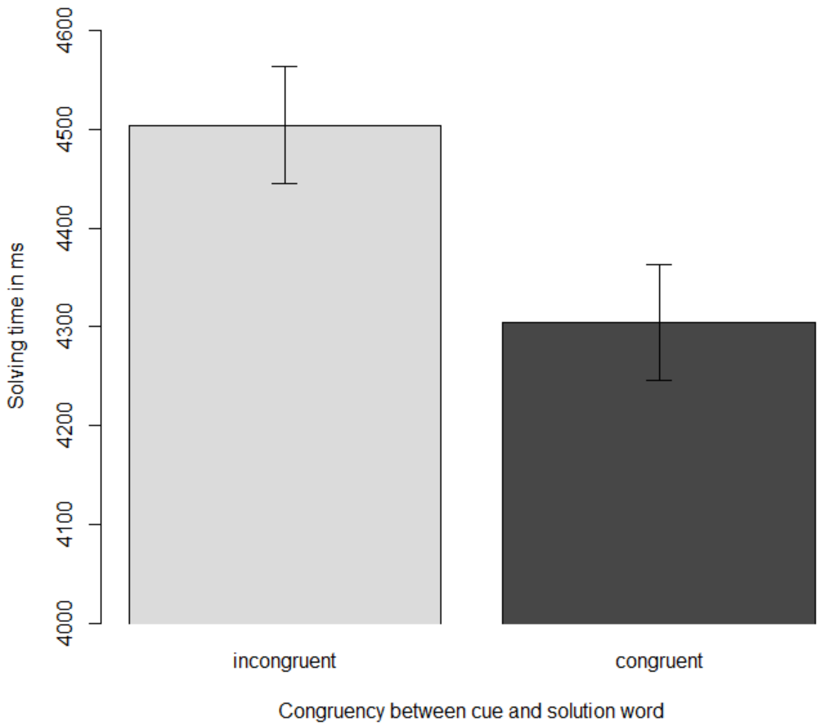


Figure 3. Solving times for anagrams with their solution words being either associated with a positive or negative emotional valence depending on whether a picture shown before the anagram was congruent with the associated valence or incongruent with it.

2.1.3. Spatial meaning domain: Attention shifts and amodal linguistic cues

The previous studies showed that modal cues can activate meaning representations in the spatial and valence domain. These results are in line with the hybrid accounts of meaning representations. Building on the finding that a pre-activation of a meaning representation seems to be possible given enough multimodal cues, we were interested whether a similar facilitation of lexical access could also be achieved when combining a vertical shift of attention with amodal linguistic cues, namely words or sentences, instead of modal cues, namely the background pictures that were used in the third and fourth experiment of Berndt, Dudschig and Kaup (2018a). If modal and amodal cues cause the same pre-activations of meaning representations then one would expect that amodal cues facilitate comprehension processes similar to modal cues. In an unpublished study, we conducted two experiments using either adjective-noun constructions or sentences ending with these adjective-noun constructions to shed light onto research question two concerning the relative role of modal and amodal representations during comprehension.

The adjective-noun constructions used in the first study consisted of nouns which referents were associated either with an upper or with a lower vertical position in space (e.g. “airplane” and “worm” respectively). The adjectives were likewise also associated with either UP or DOWN (e.g. “flying” or “dead”). We combined these adjectives and nouns so that the associated vertical position was the same for both adjective and noun (e.g. “flying airplane” or “dead worm”) or the opposite for adjective and noun (e.g. “raised anchor” or “dead eagle”), with the resulting associated position of the combined meaning of the construction always depending on the vertical position of the adjective. Similar to the procedure used in Berndt, Dudschig and Kaup (2018a), the nouns were converted to anagrams and shown either at the top or at the bottom of the screen. Before the anagram of the noun was shown, the adjective appeared at the center of the screen (see Figure 3). The adjective was supposed to serve as a linguistic cue, limiting the number of potential solutions to the anagram and at the same time imply either an upper or lower vertical position. We wanted to replicate the finding that congruent trials would be solved faster than incongruent ones and furthermore hoped that we could distinguish whether this congruency effect would depend on the associated vertical position of the noun, or whether it would depend on the associated vertical position of the adjective, which was also and the vertical position resulting from the semantic combination of adjective and noun. The latter would mean that this effect can be altered by linguistic context and is not solely dependent on a fixed association between a noun and the typical vertical position of its referent. However, the results showed that there was no difference between any of the conditions (adjective/phrase congruency: χ^2 (1) = 3.25, p = .071; noun congruency: χ^2 (1)

< 1; both adjective/phrase and noun congruency χ^2 (2) = 3.25, p = .197; interaction between adjective/phrase and noun congruency χ^2 (3) = 3.30, p = .348), indicating that the linguistic cue did not suffice to pre-activate the corresponding conceptual representation (see Figure 4) similar to what was found in the first experiment in Berndt, Dudschig and Kaup (2018a).

In another study we used the same adjective-noun constructions, only this time they appeared at the end of a simple sentence (e.g. “Tom saw the dead worm” or “Anna pointed at the flying airplane”). This was done to further emphasize the linguistic context and encourage participants to engage in a simulation of the described situation, which should be easier for complete sentences than for adjective-noun constructions without further context. However, the results were the same and again, no difference between any of the conditions was found (all χ^2 < 1, see Figure 4).

Failing to find a facilitation of lexical access in both experiments suggests that linguistic context paired with a vertical shift of attention does not pre-activate meaning representations sufficiently. This could have several reasons: First of all, the sentences and constructions used could have been too simple or too irrelevant for participants to engage in a multimodal simulation, which may have reduced the effectiveness of the vertical shift of attention as a cue. Sentences describing more complex or unusual situation could be used in similar future experiments to test this possibility. In a similar vein, the attentional shift could have been more useful as a cue when using background pictures, since they encouraged participants to engage in a perceptual simulation of the content, while these simple sentences did not require nor encourage such a simulation.

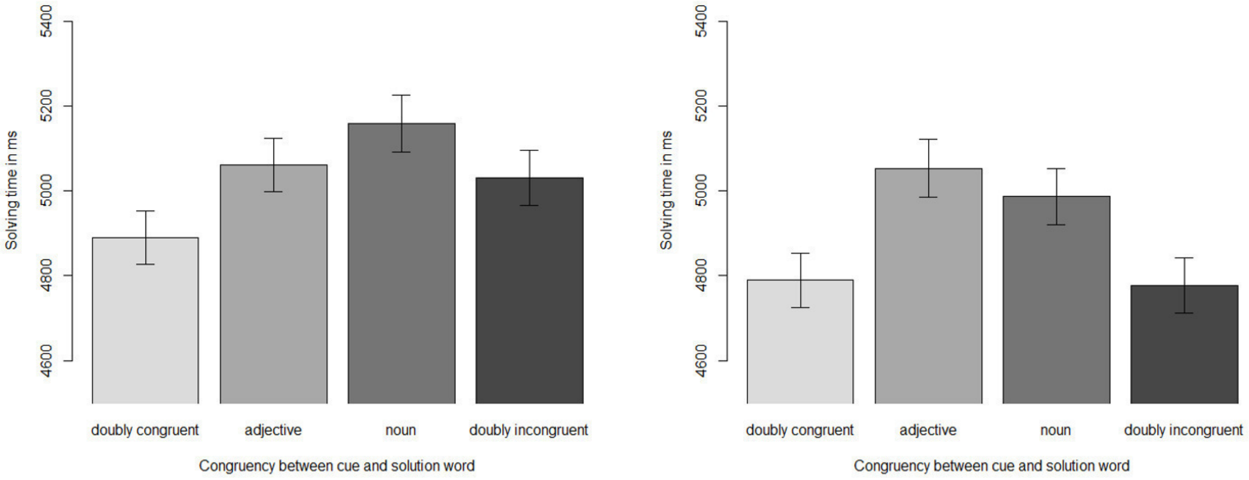


Figure 4. Solving times for anagrams solving tasks using either adjective-noun-constructions alone (left graph) or embedded in sentences (right graph). Solution words are either congruent with both the adjective and the noun, only the adjective, only the noun or neither.

2.1.4. Colour meaning domain: Modal and amodal colour cues and their relative contribution to meaning activation

After investigating the spatial and the valence meaning domain in the previous experiments - where a vertical shift of attention proved to be too subtle a cue to facilitate lexical access without additional modal cues - we decided to move to the colour domain as the relevant property in another set of experiments. Colour is an important property for object recognition (Bramao, Reis, Peterson & Faisca, 2010) and it has been shown that the typical or implied colour of an object is activated during language comprehension (Zwaan & Pecher, 2012 and Mannaert, Dijkstra & Zwaan, 2017). In Berndt, Dudschig and Kaup (2018b), we therefore decided to test and directly compare the influence of both modal and amodal color cues on subsequent anagram solving times. We expected that lexical access of words the referents of which are associated with a certain colour would be facilitated after a matching colour cue. If both colour and linguistic cues speed up anagram solving times in a similar manner, we aimed to test whether one type of cue predominates the other cue if presented simultaneously.

In the first experiment, one of four colour words (“red”, “brown”, “green” or “yellow”) was presented in the centre of the screen, followed by an anagram of a noun typically associated with one of these colours (e.g. “strawberry”, “chocolate”, “cucumber” or “banana” respectively). Indeed, a congruency effect was found, with faster solving times for anagrams following a matching colour cue ($\chi 2(1) = 7.34$, $p = .007$, $\beta = -280.62$, $t = -2.77$). The second experiment consisted of the same stimuli, only this time the color cues were rectangular colour patches and were thus classified as modal cues. The results were the same as in the first experiment: Congruent trials were solved significantly faster than incongruent ones ($\chi 2(1) = 4.27$, $p = .039$, $\beta = -218.01$, $t = -2.11$). In the third experiment, we combined both modal and amodal cues by presenting a colour word inside a rectangular colour patch. This lead to four different types of congruency: The solution word was primed by both the colour patch and the colour word (e.g. the word “green” inside a green rectangle when solving an anagram for “cucumber”), by only either the colour patch (e.g. “red” inside a green rectangle) or the colour word (e.g. “green” inside a red rectangle) or by neither of the cues (e.g. “red” inside a yellow rectangle). This allowed us to test whether either the modal or amodal cue would dominate over the other.

Interestingly, the results showed that anagrams were only solved faster when both colour word and the colour patch matched and were associated with the solution word ($\chi 2(1) = 4.41$, $p = .036$, $\beta = -231.79$, $t = -2.10$). When only either the colour word or the colour patch matched the colour associated with the solution word, solving times were equally slow as when none of the colour cues matched the colour associated with solution word ($\chi 2(1) = 0.07$, $p = .933$, $\beta = -9.446$, $t = 0.084$ and $\chi 2(1) = 0.13$, $p = .720$, $\beta = -40.24$, $t = -0.36$, respectively). Thus, if using modal and amodal cues simultaneously, the matching cue seems to have no beneficial influence on anagram solving times. In contrast, the results suggest they cancel each

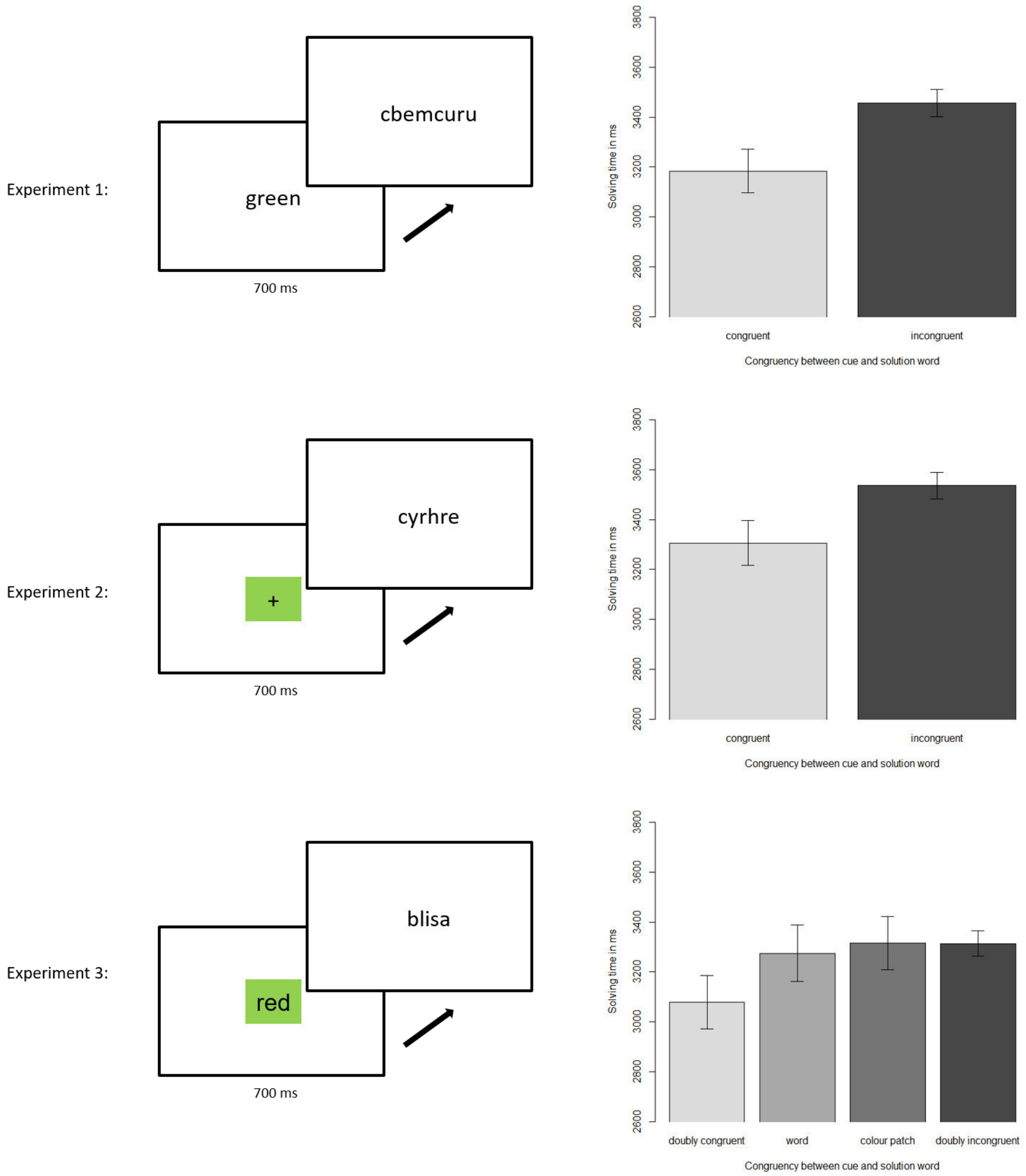


Figure 5. Solving times for anagrams of words associated with a certain colour using either amodal or modal color cues as a prime.

other out with regard to their influence on activating associated meaning representations. In summary, this result suggests that neither modal nor amodal cues outplays the other type of cue.

Taken together, the results of this set of experiments show that it is possible for experiential traces to activate meaning representations and in turn facilitate lexical access to words associated with them. Thus, these studies show that modal sensorimotor cues can also influence subsequent linguistic processes, a pre-condition for modal representations being functionally relevant for the comprehension process. Given the abstract nature of anagram solving tasks where solution words can be found by simply rearranging letters without engaging in any kind of meaning representation, the fact that multimodal cues helped participants to solve these anagrams is even more remarkable. This is why these results can be seen as evidence for the importance of multimodal experiential traces in meaning representations and the construction of meaning. Nevertheless, despite these findings supporting the pre-condition for a functional relevance of modal meaning representations – by showing that the interaction between linguistic processes and sensorimotor processed can be found in a bi-directional manner - future studies are still needed to finally exclude the option that modal representations are not at the core of the comprehension process (Mahon & Caramazza, 2008).

2.2. Hemispheric differences regarding the activation of modal meaning representations

In addition to a large number of behavioural studies of which many have been covered in the introduction, there is also a considerable number of studies investigating the neural basis of conceptual representations. It has been shown that language is processed differently in the two hemispheres in the brain (e.g. Beeman & Chiarello, 1998; Chiarello, 1991). Beeman (1998) for example assumed that fine-grained close associations are activated in the left hemisphere (LH) upon reading a word, while more coarse semantic associations are processed in the right hemisphere (RH). It has also been shown that the RH is involved in the processing of visual-spatial relations (Marsolek, Kosslyn & Squire, 1992), while the left hemisphere seems to be more involved in more abstract or more direct associations and relations like category membership (Abernathy & Coney, 1996).

This led Zwaan and Yaxley (2003a) to assume that a multimodal simulation should be confined to the RH, since spatial relations between objects are assumed to be part of such a multimodal simulation. In an earlier study (Zwaan & Yaxley 2003b), they found a match effect between the vertical arrangement of word pairs on a computer screen and their referents' canonical spatial relation. When the word "nose" was written above "mustache", participants were faster to respond in a semantic-relatedness judgement task compared to "nose" being written below "mustache", since this arrangement is supposed to be at odds with the ongoing multimodal simulation. To test for differences between both hemispheres, a divided visual field (DVF) paradigm was applied to this study. Word pairs were presented very briefly either on the left visual field (LVF) or on the right visual field (RVF). Since the presentation time was too short for participants to make a saccade towards the word pair, they were only received by the contralateral hemisphere. Using this paradigm, the expected visual field by match interaction was found: The match effect of spatial relation between objects was only present when the word pairs were presented on the LVF and thus only received by the RH, while response latencies to word pairs on the RVF were equally low regardless of match.

However, the results of this study are difficult to interpret without further controls. One issue in that study was the lack of counterbalance regarding the response side. In Zwaan and Yaxley (2003a), all of the participants had to press the J-button on the keyboard (right-handed and right-sided response) when words were semantically related and the F-button on the keyboard (left-handed and left-sided response), when the words were semantically unrelated. Since all of the experimental trials consisted of words that were semantically related, only right-sided responses were collected for the experimental trials. This could be problematic, since Simon and Rudell (1967), have shown in what is known as the Simon Effect, that responses are faster when the response side matches the presentation side of the stimulus. Thus, it is important to counterbalance response side in visual field experiments (Bourne, 2006). Therefore, we decided to replicate the Zwaan and Yaxley's (2003a) experiment with response side counterbalanced between participants (Berndt, Dudschig, Miller & Kaup, 2019).

We used German translations of the stimulus material of the original study along with the same experimental setup, we tested 96 participants in our replication attempt to achieve a power of around 80 %, which should be the case when a study is replicated with 2.5 times the number of participants compared to the original study according to Simonsohn (2015). Half of the participants responded to semantically related words by pressing the J-key, while the other half had to press the F-key in response to semantically related words. If multimodal simulations, or more precisely the representation of spatial relations between objects, are confined to the RH, we would expect to replicate the visual field by match interaction that was found by Zwaan and Yaxley (2003a), even when counterbalancing response side. If the pattern of results that lead Zwaan and Yaxley (2003a) to assume differences between our hemispheres regarding multimodal simulations was only found because of the Simon effect, we would expect to find no visual field by match interaction, but instead would expect a visual field by response side interaction.

Using linear mixed effect models, we found the data to best fit a model with a visual field by response interaction and an additional main effect of match (visual field: $\beta = 0.17$, $t = 0.02$, response side: $\beta = 47.87$, $t = 1.51$, match: $\beta = 13.71$, $t = 2.10$, visual field x response side: $\beta = -34.66$, $t = -2.65$). This means that, while there is a match effect between the spatial arrangement of words and their referents' spatial relation to each other, it does not seem to be confined to either one hemisphere (Figure 6). Instead, the Simon effect could have led to differences between both visual fields in the original study, since word pairs presented to the RVF were congruent with the response side und thus easier to respond to, while word pairs on the LVF were incongruent to the response side and thus were responded to more slowly. The match effect then could have led to faster responses to word pairs presented on the LVF,

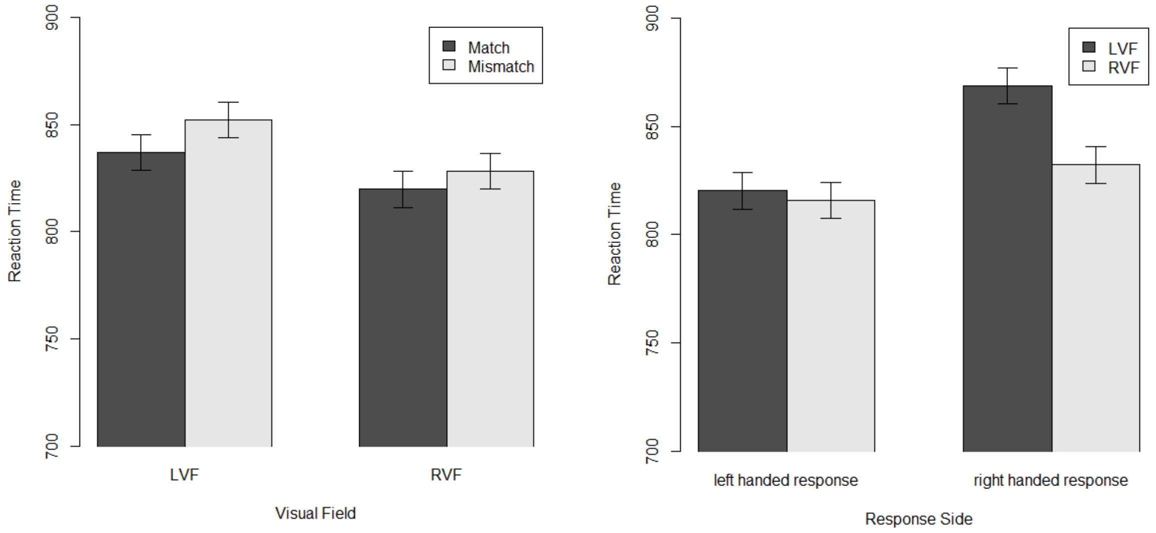


Figure 6: Summary of the results of Berndt, Dudschig, Miller and Kaup (2019): There was no interaction between Visual Field and Match (left graph) but there was a Simon-Effect (right graph). Error bars represent standard error.

bringing them on the same level as the speeded responses from the RVF that was congruent with the response side. However, this should have been reflected in a three-way interaction between match, visual field and response side, which was not present in the current data.

The findings reported here underline the importance to control for response side in a DVF paradigm, and question the assumption that multimodal experiential simulations are confined to the RH. Further experiments are needed to determine the cause of the difference in results: The first obvious difference is the use of German words instead of English words. The German translations of the original stimulus material resulted in slightly longer word length, which could have caused differences in perception of the word pairs. To account for this, another replication with English stimulus materials and English participants is thus required. Furthermore, to ensure the word pairs are actually first processed by the contralateral hemisphere, stimuli should be presented further away from the fixation point to not appear in the region of foveal overlap (see Bourne, 2006).

3. GENERAL DISCUSSION

The present thesis originated on the background of hybrid models of cognition and aimed at investigating the role of modal and amodal representations for the language system. Thereby the following core questions were in the focus: (1) Are modal representations functional for comprehension? (2) Do modal or amodal representations play a predominant role? (3) Are hemispheric representation specificities one key to differences in representational format as postulated by Zwaan and Yaxley (2003a)?

With regard to the first question – the functional role of modal representations – the studies conducted in this thesis persued the way of investigating the influence of modal cues on language processes. Therefore, anagram solving paradigms were implemented. In several experiments investigating meaning aspects in the spatial, valence and colour domain, influences of modal cues on anagram solving tasks were indeed observed. Across these domains the amount of modal aspects required to activate the relevant concepts differs substantially. In some domains a simple modal cue is sufficient to pre-activate the relevant concepts (e.g., colour and valence), whereas in other domains a combination of cues was needed (e.g,. spatial). In any case taken together these results indicate that there are influences of modal representations on language comprehension, supporting the idea that these interactions between language and sensorimotor processing can be observed in this reversed direction of influence. Findings of this type of influence are one prerequisite for the assumption that modal representations are functional for comprehension. In this vein the present results support hybrid models in which both representational formats contribute to language processes.

With regard to the second question the results are more ambiguous. In order to investigate the predominance of modal or amodal aspects for language processes we used the spatial and colour domains as fields of interest. In the spatial domain, amodal linguistic cues did not help to activate the relevant concepts, not even when combined with modal spatial cues.

In contrast, in the colour domain, amodal linguistic cues were as important as the modal colour cues. This result provided the ideal basis to investigate the predominance of one or the other type of representations format for language processes. Therefore, an experimentwhere modal and amodal cues were combined was implemented, resulting in conditions where both cues matched the target word, both mismatched the target word, as well as conflicting versions where one matched and the other mismatched the target word. Interestingly, only when both cues matched the target word a facilitatory effect on solution times was observed. As soon as one cue mismatched the colour domain of the solution word no facilitation was observed. As was argued above this suggests a symmetry concerning the importance of both types of representational formats as suggested by hybrid models.

The final key question in this thesis addressed whether there are hemispheric differences in regard to what type of representational format dominated. As a starting point a replication of a previous study pointing to such differences was conducted but with the addition of manipulating response side (Berndt, Dudschig, Miller & Kaup, 2019). In clear contrast to the original study, no hemispheric differences were observed. Rather an influence of modal representations was observed independent of hemisphere. Thus, hemispheric differences do not seem to play a key role in explaining predominances of modal and amodal representational aspects.

Taken together, the current results speak in favour of a hybrid account of cognition. Hybrid accounts of cognition (e.g. Binder & Desai, 2011; Meteyard, Cuadrado, Bahrami & Vigliocco, 2012 and Zwaan 2014 among others) assume that concepts are stored in a large-scale semantic network spanning most parts of the brain and involve both multimodal features extracted from past experiences as well as more abstract features that can be derived from multimodal information through progressive abstraction. The results of the present thesis particularly speak in favour of hybrid accounts of cognition in which the two different representational formats are functionally relevant and of equivalent importance.

3.1. Meaning features, concepts and word meaning: An integrative sketch of the results

In this dissertation several studies were discussed that in one way or the other tap into those parts of the cognitive architecture that are related to meaning, namely meaning features, lexical representations, concepts and conceptual representations. These terms are not always used unambigiously in the literature and one further difficulty arises from the fact, that it is often not completely clear which manipulation affects which aspects of this architecture and whether different tasks differ in this respect. In the following I will present an attempt to illustrate the current tasks and findings in integrated sketch of a semantic knowledge network.

Kelter and Kaup (2012) illustrated the distinction between concepts and word meanings and provided an overview over the research on concepts, conceptual representations and their connection to word forms. *Concepts* are supposedly stored in long term memory and provide information about categories and the properties of entities belonging to those categories by means of learned feature associations. A given entity can be compared to this concept and the degree of membership to the corresponding category can be judged as a function of similarity between the properties of the entity and those stored in the concept. It is noteworthy that there is an ongoing debate about the question whether concepts are abstracted from past experiences with its exemplars resulting in a prototype of a given category, called the prototype view, or if conceptual knowledge is accumulated by storing a set of exemplars of a given category, as described above. *Conceptual representations* on the other hand are representations consisting of specific activated features of concepts in working memory. These conceptual representations are not the same thing as concepts, because they are not what is stored in long-term memory but instead depend on the context and the task at hand. Furthermore, these conceptual representations are very flexible, allowing for productive combinations and propositional representations involving more than one conceptual representation. Finally, *lexical* *representations* are what constitute the lexical meaning of a given word. The equivalence view suggests that these lexical representations themselves are concepts (Murphy, 2002), however there are apparent difficulties with this view. For example, it cannot explain the existence of concepts in animals (for a review see Lazareva & Wasserman, 2008) or pre-verbal infants (for a review see Rakison & Yermolayeva, 2010). Thus, Kelter and Kaup (2012) instead assume that lexical representations, just like conceptual representations, consist of specific features that are activated in working memory upon hearing or reading a word. Applying these definitions, a semantic network was sketched by Kelter and Kaup (2012), which was based on a model by Ursino, Cuppini and Magosso (2010). They described a semantic network that consists of a large number of cognitive units representing atomic features called *microfeatures* which are interconnected with excitatory and inhibitory connections reflecting their co-occurrence or association between them in previous experiences. Concepts are basically tightly associated assemblies of microfeatures, but they are not exclusive to one single concept but instead certain overlap can exist between concepts. Word forms are also tightly connected to certain microfeatures and this pattern of connections from microfeatures to word forms is what can be described as the lexical representation. Note that it can be possible in theory that a concept and a lexical representation are identical, given all microfeatures of a concept are also connected in the same way to a word form. However, it is also possible for concepts to not be captured by a word form, taking into account that concepts can also exist without a linguistic analogue, or that more than one concept is associated with one word form, as is the case for homonyms[2].

---

[2] However, it should be noted that not all authors agree upon the differentiation between concepts and word meaning. Lupyan (2012) for example rejects this distinction. He reports findings, showing how lexical representations can alter existing concepts and even instantiate new concepts. Winawer et al. (2007) for example have shown that Russian speakers, who have two different words for the colour blue, one for dark blue and another for a lighter blue, have an advantage over English speakers in discriminating different shades of blue, when these shades of blue were falling into the different labelled categories that the Russian speakers possess. This shows that a lexical representation can directly influence cognition even on a very low level such as perception and discrimination, indicating that the categories instantiated by

In an attempt to apply this sketch of a semantic knowledge network to the results described in the previous section, we have to move one step away from microfeatures and look at smaller assemblies of microfeatures that instantiate distinguishable features. The feature <red> for example would consist of very basic visual microfeatures like wave length of light and hue that ultimately constitute what we know as red. In line with hybrid accounts of cognition, a distinction between amodal and multimodal features seems to be necessary, with amodal features resulting from a progressive abstraction of multimodal features (Binder & Desai, 2011). This would result in a knowledge network as seen in Figure 6, which will be used to outline a possible interpretation of the results reported previously.

In Berndt, Dudschig and Kaup (2018a) anagrams were solved faster when anagrams appeared at a position of the screen that matched vertical position associated with the referent of the solution word if and only if a background picture matching the theme of the solution words was presented. Attention to either the upper half of the screen could be seen as an abstract feature that is tightly connected to the features <up> (A5 in Figure 6 for example), which in turn are part of the lexical representations of "eagle" (F_W in Figure 6). But since these features are only a minor part of the lexical representations and furthermore are also connected to different lexical representations (like F_X), this feature alone did not sufficiently activate the corresponding word form F_W. Adding a background picture supposedly activated more perceptual features, for example M4, thus increasing the activation of F_W. Lastly, the letters of the anagram supposedly served as additional abstract features, further increasing the activation of F_W and inhibiting the activation of F_X. In mismatching trials, the <up> or <down> feature inhibited activation of the correct solution instead of activating it, leading to slower response times.

---

the lexical representations created two concepts of the colour blue instead of only one, in turn favouring the view that lexical representations are equal to concepts.

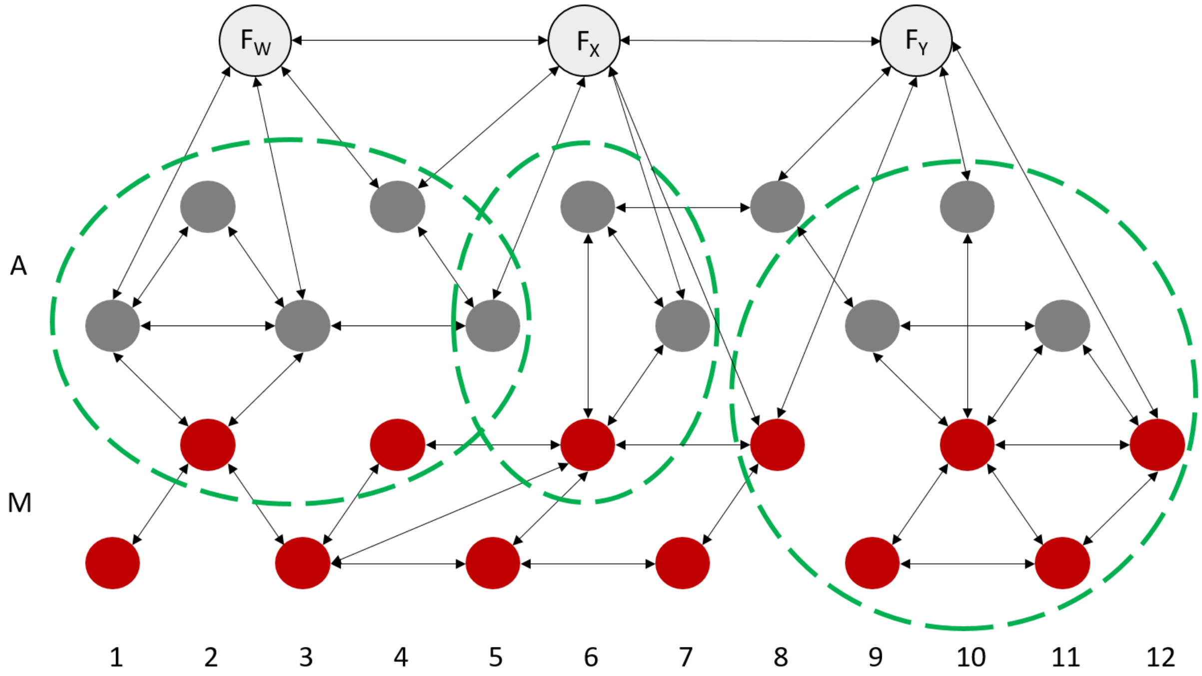


Figure 6. A sketch of a semantic knowledge network. The red nodes represent modal perceptual features that can progressively be abstracted into amodal features, depicted as dark grey nodes. Light grey nodes represent word forms that are interconnected with each other as well as an assembly of features instantiating the lexical representation of the word form. Dashed green lines indicate assemblies of features forming concepts.

A similar explanation is applicable to the results obtained in Berndt, Dudschig and Kaup (2018b) and when using pictures with emotional valence to prime positive or negative solution words. In contrast to Berndt, Dudschig and Kaup (2018a) where activation of multiple features was necessary to obtain facilitation effects on anagram solving, either a single colour patch or a colour word shown prior to anagram presentation, sufficed to facilitate anagram solving if the solution word’s referent was associated with the given colour cue. This can be explained by a more prominent role of colour in object recognition (Bramão, Reis, Peterson & Faísca, 2011), or applying it to the knowledge network, stronger connections between the colour feature and the solution word. More interestingly, there was no difference in the match effect between modal cues in the form of coloured rectangles and amodal cues, namely colour words. This indicates, that the word form of the solution word was activated equally by both types of cues. To apply this idea to the semantic network model sketched in Figure 6, assume the word form of the solution word “cucumber” is F_X, the word form of the word “green” is F_Y and the feature <green> is the node M9. Upon perceiving both a green colour patch as well as the colour word “green” inside the rectangle, F_Y and M9 are activated and in turn activate nearby associated nodes, resulting in a sufficient activation of F_X for conscious access. If however the colour word *“green”* was presented inside a rectangle coloured red, the microfeature <red> would be activated instead of M9, possibly exerting an inhibitory activation on the colour word *“green”*. Showing the colour word “red” inside a green rectangle, would activate the word form for *“green”* and the microfeature <red>, inhibiting each other in the same way. This could lead to the observed effect of the match effect being confined to trials where both cues activate the same colour concept. It would be unlikely that spreading activation of solely abstract concepts or lexical representations onto multimodal systems and then back again to these abstract concepts, as argued by Mahon and Caramazza (2008) is responsible for this pattern of results. If that would have been the case, a stronger influence on the match effect by colour words compared to coloured rectangles would have been expected. The colour words should have activated the corresponding colour concept directly, while the perception of conflicting modal cues would have activated the corresponding colour concept indirectly, with activation spreading from sensory systems to abstract conceptual systems leading to a delayed activation of the conflicting colour concept. Thus, the results reported by Berndt, Dudschig and Kaup (2018b) suggest that modal and amodal microfeatures seem to influence the activation of concepts and word forms in a similar way, rendering a functional distinction between amodal and modal representations or microfeatures questionable.

Interestingly, a linguistic context did not seem to sufficiently activate the word form of the solution word or inhibit irrelevant word forms when paired with a vertical shift of attention. This could reflect two different possible modes of processing linguistic stimuli. Presenting background pictures before the anagram could have led to participants engaging in a perceptual simulation of subsequently presented words, or in this case anagrams, which led to a facilitation of lexical access when the positional cue further matched the solution word. However, reading a simple sentence or an adjective may have discouraged participants from engaging in such a simulation and instead activating a shallower linguistic processing mode, where linguistic associations are more dominant than perceptual features. This may explain why the positional cue did have no effect, since it would not play a role in linguistic associations. This is in line with Louwerse and Jeuniaux (2010), finding stronger involvement of linguistic associations in semantic relatedness judgment tasks and when using only linguistic material, while embodied representations seem to have a bigger influence in iconicity judgment tasks or when pictures are involved in the task.

In Berndt, Dudschig, Miller and Kaup (2019) laterality differences in the simulation of spatial relations between objects was investigated. Previous findings show a dominance of the left hemisphere in general language processing (e.g. Beeman & Chiarello, 1998; Chiarello, 1991) and laterality differences in the activation of semantically associated concepts: While more abstract associations, like category membership are supposed to be processed in the LH, more distant semantic associations seem to be processed in the RH (Chiarello, 1998; Beeman, Bowden & Gernsbacher, 2000). In contrast to Zwaan and Yaxley (2003a) we were not able to replicate the finding that a match effect of spatial relation was confined to the RH, as predicted by the accounts of laterality differences mentioned previously, but instead we only found a general match effect, independent of hemisphere. While it is in principle possible to assume that word forms (F_W, F_X, F_Y in Figure 6) or either modal or amodal features (row M or row A in Figure 6) are stored and processed mainly or even exclusively in either one hemisphere, it is more likely that at they are spread across a large scale network including both hemispheres (Binder & Desai, 2011) and are not ordered as neatly across different hemispheres as was assumed. This would be in line with the assumption that modal and amodal representations are tightly interconnected, with the latter having evolved from perceptual features by progressive abstraction.

3.2. Conclusion

To conclude, the results of the experiments reported in this dissertation are in line with hybrid accounts of cognition. In particular meaning seems to be stored in large-scale networks spanning many different brain areas and both hemispheres (Berndt, Dudschig, Miller & Kaup, 2019). Depending on task and context a meaning representation containing currently relevant features, either amodal symbolic ones or multimodal, perceptual experiential traces, seems to be activated in our working memory. This meaning representation is neither rooted exclusively in amodal, symbolic systems, nor is it solely based on perceptual features. Instead modal and amodal contextual cues seem to activate related meaning representations equally strong (Berndt, Dudschig & Kaup 2018b). Further research is needed to address open questions regarding the interaction of amodal and multimodal representations and their respective roles for different cognitive processes.

4. SUMMARY (GERMAN)

In dieser Dissertation wurde der Einfluss verschiedener Kontexte und Aufgaben auf die Beschaffenheit mentaler Repräsentationen untersucht. Insbesondere ging es darum herauszufinden die Rolle multimodale Repräsentationen, die direkt mit eigenen Erfahrungen und Sinnesorganen verknüpft sind, genauer zu betrachten. Viele Studien konnten bereits zeigen, dass solche multimodale Repräsentation beim Sprachverstehen aktiviert werden und beispielsweise einen Einfluss auf motorische Reaktionen haben. Hierbei ist allerdings wenig erforscht ob die Aktivierung von multimodalen Hinweisreizen umgekehrt auch die Aktivierung damit assoziierter Wörter begünstigt. Um dies zu untersuchen wurden Anagrammlöseaufgaben verwendet, bei denen die Probanden vor den Anagrammen modale Hinsweisreize sahen, die entweder zum Lösungswort passend waren oder nicht. In Berndt, Dudschig und Kaup (2018a) wurden zunächst Nomen untersucht die entweder mit dem Himmel oder mit dem Meer assoziiert waren (z.B. “Möwe” und “Delfin”) und es zeigte sich, dass diese schneller gelöst wurden wenn sie an der passenden Position auf dem Bildschirm zu sehen waren (Lösungswort mit Himmel assoziiert oben auf dem Bildschirm und Lösungswort mit Meer assoziiert unten auf dem Bildschirm), sofern zusätzlich ein Hintergrundbild präsentiert wurde, das zum Thema passend war (Bild eines Segelbootes oder Bild des Meeres mit Horizont). Dieser Befund konnte auf Wörter erweitert werden die mit positiven oder negativen Emotionen assoziiert sind. Hierzu wurden Probanden Bilder mit positiver und negativer Valenz gezeigt, die außer der assoziierten Emotion nichts mit dem Lösungswort zu tun hatten, was dazu führte, dass Anagramme schneller gelöst wurden, wenn die assoziierte Emotion von Bild und Wort passend waren. In einer weiteren Studie wurde untersucht ob modale Hinweisreize in Form von Bildern für diese Kongruenzeffekte notwendig waren oder ob ein ähnlicher Effekt auch mit linguistischen Hinweisreizen gefunden werden kann. Die linguistischen Hinweisreize waren in diesem Fall Adjektive beziehungsweise Sätze die mit diesen Adjektiven endeten, die zum Lösungswort passten (z.B. “Fritz sah den fliegenden” gefolgt vom Anagramm VLGOE, das das Lösungswort “Vogel” ergibt). In einem letzten Anagrammexperiment wurden modale und linguistische Hinweisreize direkt miteinander verglichen. Die Lösungswörter waren hierebei Wörter die stark mit einer bestimmten Farbe assoziiert werden (z.B. “Kirsche”). Als modale Hinweisreize dienten farbige Rechtecke während linguistische Hinweisreize die entsprechenden Farbwörter waren (z.B. “rot”). Nachdem gezeigt wurde, dass diese beiden Hinweisreize in ähnlichem Maße das Lösen von Anagrammen von mit den Farben assoziierten Wörter erleichtern, wurden beide Hinweisreize gleichzeitig verwendet (z.B. “grün” in einem roten Rechteck vor einem Anagramm zu “Gurke”). Hierbei zeigte sich, dass weder modale noch linguistische Hinweisreize übereinander dominierten. Stattdessen wurden Anagramme nur dann schneller gelöst, wenn beide Hinweisreize zum Lösungswort passten. Nachdem mit diesen Anagrammstudien gezeigt werden konnte, dass sowohl multimodale als auch abstrakte Repräsentationen je nach Kontext und Aufgabe aktiv sein können, wurde letztlich noch untersucht ob diese beiden Arten von Repräsentationen verschiedenen Hirnhälften zuzuordnen sind. Einen Hinweis darauf fanden Zwaan und Yaxley (2003a), die zeigen konnten, dass visuell-räumliche Relationen von Wortpaaren nur dann einen Einfluss auf Reaktionszeiten bei einer Aufgabe zur Bestimmung semantischer Ähnlichkeit hatten, wenn sie in der rechten Hemisphäre verarbeitet wurden. Die Wortpaare wurden sehr kurz entweder auf der linken oder rechten Seite des Bildschirms präsentiert und standen entweder in der der bildlichen Vorstellung entsprechenden richtigen oder falschen Anordnung übereinander (“Nase” über “Mund” beziehungsweise “Mund” über “Nase”). Diese Studie wurde repliziert, wobei zusätzlich die Antwortseite ausbalanciert wurde, da in der Originalstudie alle Reaktionen auf die entscheidenden Wortpaare mit der rechten Hand erfolgte, was zu schnelleren Reaktionen auf Stimuli die auf der rechten Seite präsentiert werden führen sollte. Im Gegensatz zur Originalstudie wurde kein Zusammenhang zwischen der vertikalen Anordnung der Wortpaare und der Präsentationsseite gefunden. Stattdessen zeigte sich nur ein Zusammenhang zwischen

Präsentationsseite und Antwortseite. Daraus lässt sich folgern, dass multimodale Repräsentationen nicht auf die rechte Hemisphäre beschränkt sind, sondern zusammen mit amodalen Repräsentationen in Form von weitverzweigten Netzwerken über das gesamte Gehirn verteilt sind.

Zusammenfassend sprechen die hier berichteten Ergebnisse für ein Hybridmodell der Kognition, bei dem übergeordnete konzeptuelle Repräsentationen den Kern semantischer Bedeutung bilden und gleichermaßen abhängig vom jeweiligen Kontext von modalen und amodalen Reizen aktiviert werden können.

5. REFERENCES

Abernathy, M., & Coney, J. (1996). Semantic category priming in the left hemisphere. *Neuropsychologia*, *34*, 339-350. http://dx.doi.org/10.1016/0028-3932(95)00144-1

Areshenkoff, C. N., Bub, D. N., & Masson, M. E. (2017). Task-dependent motor representations evoked by spatial words: Implications for embodied accounts of word meaning. *Journal of Memory and Language*, *92*, 158-169. https://doi.org/10.1016/j.jml.2016.06.006

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. https://doi.org/10.1146/annurev.psych.59.103006.093639

Beeman M (1998) Coarse semantic coding and discourse comprehension. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255–284). Mahwah, NJ: Lawrence Erlbaum Associates.

Beeman, M. J., Bowden, E. M., & Gernsbacher, M. A. (2000). Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language*, *71*, 310-336. https://doi.org/10.1006/brln.1999.2268

Beeman, M. J., & Chiarello, C. (1998). Complementary right-and left-hemisphere language comprehension. *Current Directions in Psychological Science*, *7*, 2-8. https://doi.org/10.1111/1467-8721.ep11521805

Berndt, E., Dudschig, C., & Kaup, B. (2018a). Activating concepts by activating experiential traces: investigations with a series of anagram solution tasks. *Quarterly Journal of Experimental Psychology*, *71*, 483–498. https://doi.org/10.1080/17470218.2016.1261913

Berndt, E., Dudschig, C., & Kaup, B. (2018b). Green as a cbemcuru: modal as well as amodal color cues can help to solve anagrams. *Psychological Research*. https://doi.org/10.1007/s00426-018-1055-y

Berndt, E., Dudschig, C., Miller, J., & Kaup, B. (2019). A replication attempt of hemispheric differences in semantic-relatedness judgments (Zwaan & Yaxley, 2003). *Acta Psychologica*, *198*, 102871. https://doi.org/10.1016/j.actpsy.2019.102871

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*, 527-536. https://doi.org/10.1016/j.tics.2011.10.001

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*, 1-28. https://doi.org/10.1016/S0010-0277(99)00073-6

Bourne, V. J. (2006). The divided visual field paradigm: Methodological considerations. *Laterality*, *11*, 373-393. https://doi.org/10.1080/13576500600633982

Bramão, I., Faísca, L., Forkstam, C., Reis, A., & Petersson, K. M. (2010). Cortical Brain Regions Associated with Color Processing: An FMRI study. *The Open Neuroimaging Journal*, *4*, 164-173. http://dx.doi.org/10.2174/1874440001004010164

Bub, D. N., Masson, M. E., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*, 27-58. https://doi.org/10.1016/j.cognition.2006.12.010

Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right-and left-handers. *Journal of Experimental Psychology: General*, *138*, 351-367. http://dx.doi.org/10.1037/a0015854

Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, *2*, 79-116. https://doi.org/10.1515/langcog.2010.004

Chiarello, C. (1991). Interpretation of word meanings by the cerebral hemispheres: one is not enough. In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 251–278). Hillsdale, NJ: Erlbaum.

Chiarello, C. (1998). On Codes of Meaning and the Meaning of Codes, Semantic Access and Retrieval within and between hemispheres. In M. Beeman & C. Chiarello (eds.) *Right hemisphere language communication, Perspectives from Cognitive Neuroscience*, (pp. 141-160). Mahwah, NJ: Erlbaum.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1975). *Reflections on Language*. New York, NY: Pantheon Books.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*, 1-15. https://doi.org/10.1017/S0140525X00001515

Cuppini, C., Ursino, M., Magosso, E., Rowland, B. A., & Stein, B. E. (2010). An emergent model of multisensory integration in superior colliculus neurons. *Frontiers in Integrative Neuroscience*, *4*, 6. https://doi.org/10.3389/fnint.2010.00006

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25-62. https://doi.org/10.1016/0010-0277(89)90005-X

de la Vega, I., De Filippis, M., Lachmair, M., Dudschig, C., & Kaup, B. (2012). Emotional valence and physical space: limits of interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 375-385. http://dx.doi.org/10.1037/a0024979

Dominowski, R. L., & Ekstrand, B. R. (1967). Direct and associative priming in anagram solving. *Journal of Experimental Psychology*, *74*, 84-86. http://dx.doi.org/10.1037/h0024509

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*, 412-431. https://doi.org/10.1016/j.cognition.2008.11.016

Dudschig, C., & Kaup, B. (2017). Is it all task-specific? The role of binary responses, verbal mediation, and saliency for eliciting language-space associations. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *43*, 259-270. https://doi.org/10.1037/xlm0000297

Eikmeier, V., Alex-Ruf, S., Maienborn, C., & Ulrich, R. (2015). How strongly linked are mental time and space along the left–right axis?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1878-1883. https://doi.org/10.1037/xlm0000129

Ellis, J. J., Glaholt, M. G., & Reingold, E. M. (2011). Eye movements reveal solution knowledge prior to insight. *Consciousness and Cognition*, *20*, 768-776. https://doi.org/10.1016/j.concog.2010.12.007

Fodor, J. A. (1975). *The language of thought.* Cambridge, MA: Harvard University Press.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3-71. https://doi.org/10.1016/0010-0277(88)90031-5

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*, 455-479. https://doi.org/10.1080/02643290442000310

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*, 905-922. https://doi.org/10.1016/j.cortex.2011.04.010

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558-565. https://doi.org/10.3758/BF03196313

Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 43, pp. 93-126). New York, NY: Elsevier Science.

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, *18*, R290-R291. https://doi.org/10.1016/j.cub.2008.02.036

Goldinger, S. D., Papesh, M. H., Barnhart, A. S., Hansen, W. A., & Hout, M. C. (2016). The poverty of embodied cognition. *Psychonomic Bulletin & Review*, *23*, 959-978. https://doi.org/10.3758/s13423-015-0860-1

Gozli, D. G., Chasteen, A. L., & Pratt, J. (2013). The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology: General*, *142*, 1028-1046. http://dx.doi.org/10.1037/a0030362

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*, 335-346. https://doi.org/10.1016/0167-2789(90)90087-6

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301-307. https://doi.org/10.1016/S0896-6273(03)00838-9

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve?. S*cience*, *298*, 1569-1579. https://doi.org/10.1126/science.298.5598.1569

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kaschak, M. P., 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. https://doi.org/10.1016/j.cognition.2004.06.005

Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, *30*, 733-744. https://dx.doi.org/10.1207/s15516709cog0000_54

Kelter, S. & Kaup, B. (2012): Conceptual knowledge, categorization, and meaning. In: Maienborn, C., Heusinger, K. von & Portner, P. (eds.): *Semantics: An International Handbook of Natural Language Meaning* (pp. 2775-2804). Berlin: Walter de Gruyter.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182. http://dx.doi.org/10.1037/0033-295X.95.2.163

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*, 109. http://dx.doi.org/10.1037/0033-295X.92.1.109

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394. http://dx.doi.org/10.1037/0033-295X.85.5.363

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, *140*, 14-34. http://dx.doi.org/10.1037/a0021446

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychonomic Bulletin & Review*, *18*(6), 1180-1188. https://doi.org/10.3758/s13423-011-0158-x

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual.* Gainesville, FL: University of Florida.

Lazareva, O. F., & Wasserman, E. A. (2008). Category learning and concepts in animals. In: R. Menzel (ed.). *Learning and Memory: A comprehensive Reference, vol. 1: Learning Theory and Behavior*. Oxford: Elsevier, 197-226.

Leshinskaya, A., & Caramazza, A. (2016). For a cognitive neuroscience of concepts: Moving beyond the grounding issue. *Psychonomic Bulletin & Review*, *23*, 991-1001. https://doi.org/10.3758/s13423-015-0870-z

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*, 96-104. https://doi.org/10.1016/j.cognition.2009.09.002

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology*, *3*, 54. https://doi.org/10.3389/fpsyg.2012.00054

Mahon, B. Z. (2015). What is embodied about cognition?. *Language, Cognition and Neuroscience*, *30*, 420-429. https://doi.org/10.1080/23273798.2014.987791

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*, 59-70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Mannaert, L. N. H., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation?. *Memory & Cognition*, *45*, 974-982. https://doi.org/10.3758/s13421-017-0708-1

Marsolek, C. J., Kosslyn, S. M., & Squire, L. R. (1992). Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 492. http://dx.doi.org/10.1037/0278-7393.18.3.492

Masson, M. E., & Loftus, G. R. (2003). Using Confidence Intervals for Graphically Based Data Interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220. http://dx.doi.org/10.1037/h0087426

McClelland, J. L., Rumelhart, D. E., & PDP Research Group. (1986). *Parallel distributed processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science*, *15*, 243-247. https://doi.org/10.1111/j.0956-7976.2004.00659.x

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*, 788-804. https://doi.org/10.1016/j.cortex.2010.11.002

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, *18*, R732-R733. https://doi.org/10.1016/j.cub.2008.07.016

Murphy, G. L., 2002. *The Big Book of Concepts*. Cambridge, MA: The MIT Press.

Paivio, A. (1986). Mental representations: A dual-coding approach. New York, NY: Oxford University Press.

Paivio, A., & Ernest, C. H. (1971). Imagery ability and visual perception of verbal and nonverbal stimuli. *Perception & Psychophysics*, *10*, 429-432. https://doi.org/10.3758/BF03210327

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*, 976-987. https://doi.org/10.1038/nrn2277

Pecher, D., Boot, I., & Van Dantzig, S. (2011). Abstract concepts: Sensory-motor grounding, metaphors, and beyond. In B.H. Ross (Ed.), *The psychology of learning and motivation* (pp. 217-248). Amsterdam: Academic Press.

https://doi.org/10.1016/B978-0-12-385527-5.00007-3

Pecher, D., Van Dantzig, S., Boot, I., Zanolie, K., & Huber, D. E. (2010). Congruency between word position and meaning is caused by task-induced spatial attention. *Frontiers in Psychology*, *1*, 1-8. https://doi.org/10.3389/fpsyg.2010.00030

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793-797. https://doi.org/10.1111/j.1460-9568.2005.03900.x

Pylyshyn, Z. W. (1989). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Rakison, D. H., & Yermolayeva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 894-905. https://doi.org/10.1002/wcs.81

Rueschemeyer, S. A., Lindemann, O., van Rooij, D., van Dam, W., & Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology, 57*. https://doi.org/10.1027/1618-3169/a000031

Santiago, J., Lupáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512-516. https://doi.org/10.3758/BF03194099

Schuil, K. D., Smits, M., & Zwaan, R. A. (2013). Sentential context modulates the involvement of the motor cortex in action language processing: An fMRI study. *Frontiers in Human Neuroscience*, *7*, 100. https://doi.org/10.3389/fnhum.2013.00100

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417-424. https://doi.org/10.1017/S0140525X00005756

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*, 297-312. https://doi.org/10.1080/01690960600732430

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300-304. http://dx.doi.org/10.1037/h0020586

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *26*, 559-569. https://doi.org/10.1177/0956797614567341

Strozyk, J. V., Dudschig, C., & Kaup, B. (2019). Do I need to have my hands free to understand hand-related language? Investigating the functional relevance of experiential simulations. *Psychological Research*, *83*, 406-418. https://doi.org/10.1007/s00426-017-0900-8

Strozyk J. V., Leuthold, H., Miller, J. and Kaup B. (2018). Unpublished manuscript, University of Tübingen, Germany.

Ulrich, R., Eikmeier, V., de la Vega, I., Fernández, S. R., Alex-Ruf, S., & Maienborn, C. (2012). With the past behind and the future ahead: Back-to-front representation of past and future sentences. *Memory & Cognition*, *40*, 483-495. https://doi.org/10.3758/s13421-011-0162-4

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(, 7780-7785. https://doi.org/10.1073/pnas.0701644104

Wolter, S., Dudschig, C., de la Vega, I., & Kaup, B. (2015). Musical metaphors: Evidence for a spatial grounding of non-literal sentences describing auditory events. *Acta Psychologica*, *156*, 126-135. https://doi.org/10.1016/j.actpsy.2014.09.006

Yee, E., Chrysikou, E. G., Hoffman, E., & Thompson-Schill, S. L. (2013). Manual experience shapes object representations. *Psychological Science*, *24*, 909-919. https://doi.org/10.1177/0956797612464658

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, *18*, 229-234. https://doi.org/10.1016/j.tics.2014.02.008

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. Zwaan (ed.) *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought* (p. 224-245). New York, NY: Cambridge University Press.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS one*, *7*, e51382. https://doi.org/10.1371/journal.pone.0051382

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168-171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. A., & Yaxley, R. H. (2003a). Hemispheric differences in semantic-relatedness judgments. *Cognition*, *87*, B79-B86. https://doi.org/10.1016/S0010-0277(02)00235-4

Zwaan, R. A., & Yaxley, R. H. (2003b). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*, 954-958. https://doi.org/10.3758/BF03196557

6. APPENDIX

A. Study 1:

Copyright Notice

Berndt, E., Dudschig, C., & Kaup, B. (2018). Activating concepts by activating experiential traces: investigations with a series of anagram solution tasks. Quarterly Journal of Experimental Psychology, 71, 483–498. https://doi.org/10.1080/17470218.2016.1261913

Original Article

QJEP

Quarterly Journal of Experimental Psychology
2018, Vol. 71(2) 483–498

DOI: 10.1080/17470218.2016.1261913
qjep.sagepub.com
SAGE

Activating concepts by activating experiential traces: investigations with a series of anagram solution tasks

Eduard Berndt[1], **Carolin Dudschig**[1] **and Barbara Kaup**[1]

Abstract

According to the experiential-simulations view of language, words automatically activate experiential traces that stem from the reader's interactions with their referents. Here, we focus on the corresponding influence in the opposite direction. By means of an anagram-solving task we investigated whether activating spatial experiential traces would activate the corresponding concepts, which in turn facilitates access to associated words. Participants solved anagrams of nouns associated with the ocean or the sky (e.g. *dolphin* = "*dplhion*" or *cloud* = "*cdulo*"). In six experiments we provided additional context information such as positional information (presenting the anagram at the top or the bottom of the screen), or pictorial information that either matched the ocean and sky theme or not, or both positional and pictorial information. Anagrams were solved faster when the position of the anagram was congruent with the location of the noun's referent in the real world, but only when presented on the background of an ocean-sky picture. Thus, activating experiential traces indeed seems to activate related concepts but positional information alone is not enough to find facilitation in an anagram solving task. Rather what is needed is a whole set of traces that sufficiently narrow down the number of related concepts.

Keywords

Anagrams; Experiential traces; Grounded cognition; Meaning representation; Problem solving

Received: 15 January 2016; accepted: 27 October 2016

The traditional view on human cognition assumes that concepts are represented in an amodal, arbitrary, and abstract fashion and that these representations are isolated from other modalities in the human brain (e.g., Fodor, 1975). This view has been challenged by embodied theories of human cognition (Barsalou, 1999), supposing that concepts are represented by multimodal perceptual symbols and therefore are grounded in action, perception, and emotion. In particular, it is assumed that concepts are directly associated with experiential traces that stem from experiencing the corresponding objects, situations, and events. It is furthermore assumed that these traces are associated with the linguistic labels used to refer to the corresponding entities (Zwaan & Madden, 2005). In other words, according to this account of human cognition, words should activate the experiential traces associated with their referents and vice versa. Thus, upon hearing or reading a word like "*sun*", the experiential traces—for example, stemming from seeing (e.g., shape or typical location) or feeling the sun (e.g., warmth)—should become re-activated and build the core of the comprehension processes.

There have been many studies providing evidence for this embodied view of cognitive processing. In one experiment by Zwaan, Stanfield, and Yaxley (2002), participants read sentences like "*He saw the eagle in the sky*" or "*He saw the eagle in the nest*", followed by a picture. Responses were faster when the picture matched the sentence (an eagle with outstretched wings following the sentence "*He saw the eagle in the sky*" or with its wings drawn in after the sentence "*He saw the eagle in the nest*") than when it did not. This supports the idea that a multimodal simulation of the situation described in the sentence is activated, either supporting or interfering with the picture shown after the sentence. Glenberg and Kaschak (2002) also found compatibility effects between motor responses away

---

[1] University of Tübingen, Psychologisches Institut, Tübingen, Germany

Corresponding author:
Eduard Berndt, University of Tübingen, Psychologisches Institut, Schleichstrasse 4, Tübingen 72076, Germany.
Email: eduard.berndt@uni-tuebingen.de

from or towards one's body and sentences that describe similar movements (*"He closed the drawer"* or "*He opened the drawer"*). These compatibility effects have also been found when using single words instead of short sentences with a variety of tasks: Meier and Robinson (2004) found that there is an association between word valence and vertical space by observing that positive words (e.g., "*hero")* were evaluated faster when shown at the top of a computer screen than when they were shown at the bottom of the screen, while it was the other way round for negative words (e.g., "*liar"*) (see also: Dudschig, de la Vega, & Kaup, 2015). Similar results were obtained by Lachmair, Dudschig, De Filippis, de la Vega, and Kaup (2011) and Dudschig, Lachmair, de la Vega, De Filippis, and Kaup (2012), who have shown an association between words referring to entities with a typical location in space—or verbs describing vertical movements—and vertical motor responses. Participants responded faster after seeing a word associated with an upper vertical space (e.g., "*airplane" or "rise"*, respectively) compared to a downward response after the same word, while the opposite was true for words associated with lower vertical space (e.g., "*root" or "sink"*). Interestingly, vertical eye movement responses were also influenced by preceding location words (Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013). Further evidence for a strong connection between language and action comes from a study by Bub, Masson, and Cree (2008) showing that words denoting objects that are associated with a certain grasp type automatically evoke the corresponding grasp in the same way that pictures of those objects did. Additionally, a study by Pulvermüller, Hauk, Nikulin, and Ilmoniemi (2005) provides evidence for a functional link between motor areas and language, showing that application of transcranial magnetic stimulation (TMS) to arm- or leg-related motor areas in the brain facilitated the recognition of words associated with that motor area (e.g., "*pick"* or "*kick"*). Taken together, all these findings suggest an automatic activation of a multimodal concept representation upon reading single words or sentences.

Despite all of this evidence pointing towards such a multimodal representation there are still valid concerns: Some of these effects have been shown to also be explainable solely by language association statistics (see Goodhew, McGaw, & Kidd, 2014; Louwerse & Jeuniaux, 2010). It has also been shown that most of the effects described above are task or context dependent (see: Areshenkoff, Bub, & Masson, 2017; Brookshire, Casasanto, & Ivry, 2010; Dudschig & Kaup, 2016; Lebois, Wilson-Mendenhall, & Barsalou, 2015; Schuil, Smits, & Zwaan, 2013) raising the question if and when a supposed multimodal representation is activated and what the role of this representation is. To illustrate, in a study by Šetić and Domijan (2007) it was found that words for flying animals were judged faster when they appeared at the top of the screen than when they appeared at the bottom of the screen, while words for non-flying animals were judged faster when shown at the bottom of the screen than at the top of the screen. This result seems to speak in favour of embodied models of cognitive processing. However, Pecher, Van Dantzig, Boot, Zanolie, and Huber (2010) conducted a rather similar experiment, with words related to the ocean or the sky being shown at the top or at the bottom of the screen using two semantic decision tasks. Participants decided whether the item described by the word could be found in the sky or in the ocean. They found an interaction between task type and position on the screen, with faster responses for the ocean task when a word was shown at the bottom of the screen than at the top of the screen and faster responses for the sky task when a word appeared at the top of the screen than when it appeared at the bottom. To make matters even more complicated, Estes, Verges, and Barsalou (2008) found inhibition effects in very similar studies. They argued that this inhibition happens due to an attentional shift followed by a perceptual simulation of the described object, which hinders target detection in compatible locations.

In summary, the results reported across various studies investigating the activation of multimodal representations during comprehension are manifold. Despite these results being shown across various domains (perceptual and motor) and for various linguistic input (e.g., sentences, location words, valence words), there is one key question that has not yet received much attention in this literature and that we aim to address in our study: The previous studies almost solely focused on the influence of language on perceptual or motor systems during or after reading words or sentences. If single words or sentences trigger experiential traces, leading to a multimodal representation of the object or situation described, it should also be possible to activate certain words when enough experiential traces associated with their meaning are activated. In other words, whereby the literature strongly focused on whether experiential traces are activated following a linguistic input, the question whether an experiential trace activates a concept or word has received way less attention. Interestingly, this reversed way of influence is almost more crucial for the literature on embodied language processing, as it suggests that experiential traces are not solely a by-product of language processing, but in contrast have a direct influence on language processes themselves.

One possible way to analyse the influence of different perceptual cues on concept activations would be to use an anagram-solving task, a task that has mainly been used in the problem-solving literature. Anagrams are strings of letters, resulting from scrambling the letters of a word; "*krocte"*, for example, would be an anagram for the solution word "*rocket"*. Novick and Sherman (2003) reported two, supposedly qualitatively different, ways of solving anagrams: search solutions, involving a serial procedure of hypothesis testing by rearranging the letters, and pop-out solutions, where the solution seems to come into mind suddenly without awareness of an active solution attempt. They showed that both ways rely on the gradual accumulation of partial information (see also Ellis, Glaholt, & Reingold, 2011). These results suggest that anagram solving can be conceptualized either as a problem-solving task, where letters have to be rearranged until a solution word is found, or as a lexical access task, where letters serve as cues that activate the solution word (Fink & Weisberg, 1981). In line with this latter assumption, Dominowski and Ekstrand (1967) showed that solving times for anagrams can be reduced significantly by providing additional semantic cues instead of having only the rearranged letters as cues, for instance by priming with semantically associated words (*"table"* shown before an anagram of "*chair"*, for example). In the present study, we aimed at investigating whether activating non-linguistic experiential traces presumably associated with the solution words would also facilitate anagram solution times.

In our current study we decided to use mostly the same stimulus material as that also used in Pecher et al. (2010) with a similar set-up: Ocean (e.g., "*dolphin"*) or sky words (e.g., "*cloud"*) were scrambled to create anagrams, leaving the first letter in the correct position to narrow down possible solutions. These anagrams were presented at either the top or the bottom of the computer screen as a spatial cue (see, for example, Johansson & Johansson, 2014, for the influence of eye movements on memory retrieval) resulting in congruent (sky anagrams appearing at the top or ocean anagrams appearing at the bottom of the screen) and incongruent trials (sky anagrams appearing at the bottom or ocean anagrams appearing at the top of the screen). Our hypothesis was that anagram solution in congruent trials should be faster than in incongruent trials. Since it has already been shown that different contexts and different tasks can modulate such congruency effects, we conducted multiple experiments with different background pictures, either associated with the sky/ocean theme or not.

Experiment 1

Method

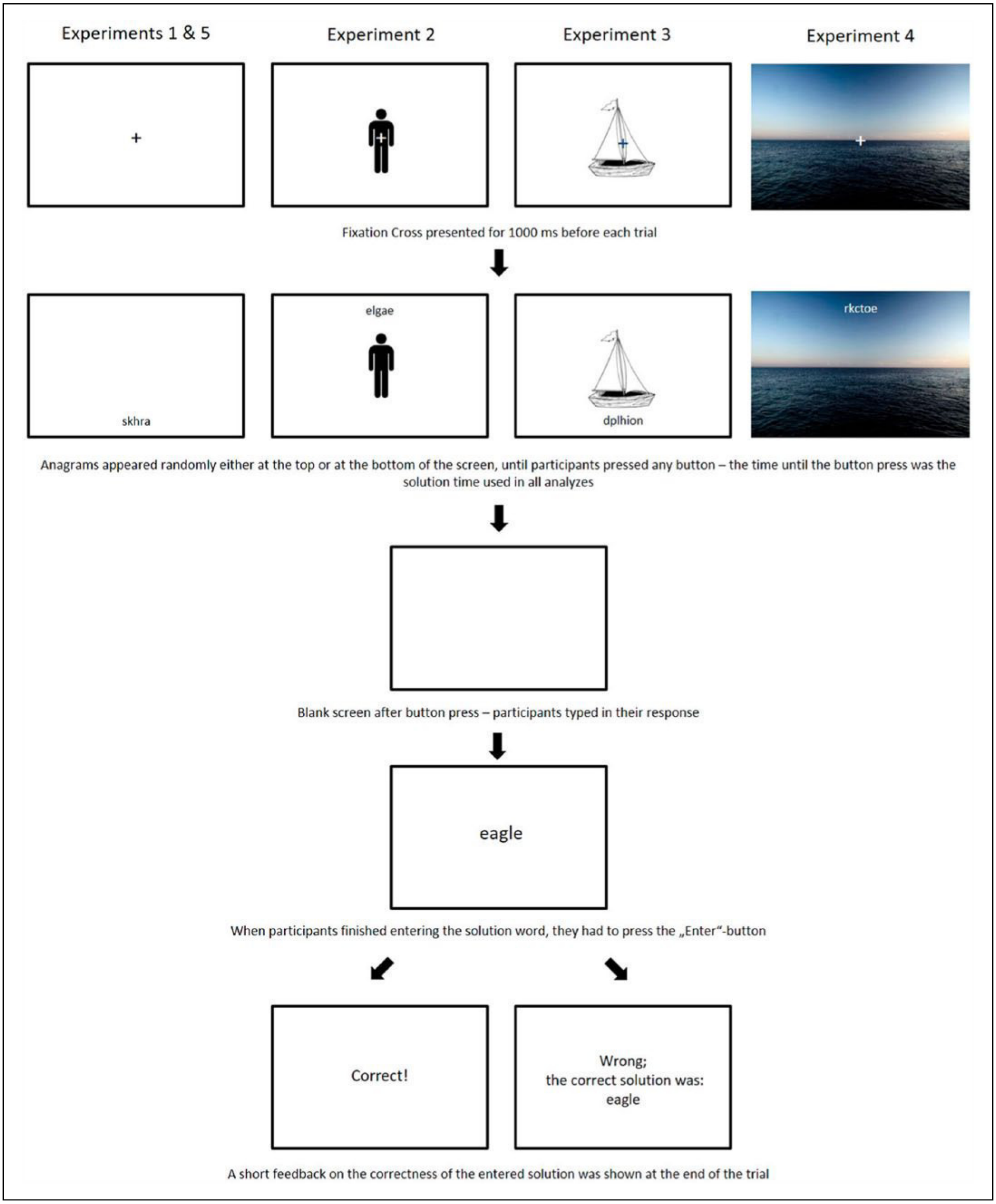
Participants. Twenty-eight German native students from the University of Tübingen, with an age ranging from 18 to 29 years (M=22.96, SD=3.16) participated in the first experiment for money or course credit. Twenty of them were female. All participants gave written informed consent.

Materials. The stimulus materials comprised 60 words in total. Fifty-one of these words were taken from the word list of Pecher et al. (2010), who had them rated regarding their association to either the sky or the ocean theme. We translated these words into German. To control for word length, nine words were added that were also associated with the sky (seven new words) or the ocean (two new words) theme. In the end, we had 30 words for each category. The word length ranged from 4 to 11 letters (M=6.67) and was matched between categories, so that an equal number of words of each length was in each category. To check for their associated vertical position, all words were rated on a scale of 1 (very low) to 7 (very high) regarding its referent's natural position in the real world. Ocean words were rated to be significantly lower than sky words ($M_{\text{ocean}}=2.51$, $M_{\text{sky}}=6.14$), $t(29)=-28.21$, $p<.001$. We did not control for frequency classes between word categories, resulting in a significant difference between both groups, $t(29)=3.62$, $p<.01$, with ocean words having higher frequency classes than sky words (frequency classes were taken from the German corpus "Wortschatz Universität Leipzig", http://wortschatz.uni-leipzig.de). However, this should not negatively influence our testing, since we are looking for an interaction between word category and presentation location (congruency) and are not looking at main effects of word category. Anagrams were generated by scrambling all letters except the first one for all of the 60 target words. Because the first letter was given, our solution words were the only sensible solutions to those anagrams. A list with all the anagrams and solution words and their corresponding English translations can be found in the Appendix.

Procedure. The participants were seated in a sound-attenuated laboratory and were given a written instruction on a Windows XP computer with a 19″ monitor, using E-Prime (Version 2.1.0). They were asked to solve the anagrams presented on the screen, with all solution words being nouns, and that they should press any key, as soon as they had come up with a solution word. Subsequently they should enter the solution word on the computer keyboard. To make sure that the participants understood everything correctly, they went through a short training block, consisting of four trials of anagrams that were not used in the experiment proper.

Each trial started with a short fixation cross in the middle of the screen, presented for 1000 ms, followed by the presentation of an anagram either at the top or at the bottom of a blank white screen. The anagram remained on the screen until a key was pressed. Then, the anagram disappeared, and the participants were able to enter the solution word, seeing what they were typing on the screen. They pressed the "Enter"-key to enter the solution. Then the feedback appeared, displaying "Correct!", if the answer was correct or "The right word would have been . . . ", if the answer was wrong. If the participants were not able to come up with a solution, they could abort the trial by pressing a key and then entering no solution word. The trial procedure is illustrated in Figure 1.

Half of the anagrams of each category were presented at the bottom of the screen, while the other half were



Fixation Cross presented for 1000 ms before each trial

Anagrams appeared randomly either at the top or at the bottom of the screen, until participants pressed any button – the time until the button press was the solution time used in all analyzes

Blank screen after button press – participants typed in their response

When participants finished entering the solution word, they had to press the „Enter"-button

A short feedback on the correctness of the entered solution was shown at the end of the trial

Figure 1. The structure of a single trial in Experiments 1–5. After the feedback slide, the next trial started. To view this figure in colour, please visit the online version of this Journal.

presented at the top of the screen. Thus, each participant saw each anagram only once at one position on the screen. There were two versions of the experiment, with the difference being that anagrams presented at the top of the screen in one version were shown at the bottom of the screen in the other version, and those presented at the bottom of the screen in one version were presented at the top of the screen. This was varied between participants, because seeing the same anagram twice would be more of a recall task than a true anagram-solving task.

Table 1. Mean solving times and percentage of correctly solved anagrams in Experiment 1.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	9436 (1428)	77.92 (3.15)	9469 (1401)	76.43 (3.19)
Sky words	8976 (1428)	78.81 (3.15)	9526 (1401)	79.76 (3.19)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).

Design. The experimental design was a 2 × 2 within-subjects design, with position on screen (bottom vs. top) and word category (ocean vs. sky) as independent variables and the time from presentation of the anagram until the key press (solving time) as the dependent variable. These two factors, however, were combined to the factor congruency, because we used linear mixed effect models and needed to reduce the complexity of the models in order for them to reach convergence.

Results and discussion

All data were analysed with the free statistics software R (Version 2.16). Since there are large individual differences in the ability to solve anagrams, and the solving times moreover strongly depend on a number of specific characteristics of the solution words, we employed a two-step procedure to eliminate outliers: First, all valid trials of each participant were converted to *z*-scores. Then solving times with a *z*-score that deviated more than 2.5 standard deviations from the mean *z*-score of the respective solution word in the respective condition were discarded (see Kaup, Lüdtke, & Zwaan, 2006). This eliminated less than 3% of the data. We analysed the data with linear mixed effect models (LMEMs), using the package lme4 for R (Bates, Mächler, Bolker, & Walker, 2015). We report chi-square values with their corresponding *p*-values for model comparisons as well as β-estimates for the congruency factor along with their corresponding *t*-value and .95 Wald confidence interval (CI). Our base model contained a fixed main effect for anagram length, random intercepts for subjects and items, and by-item random slopes for presentation location and by-subject random slopes for congruency, with congruent trials being ocean words presented at the bottom of the screen or sky words being presented at the top of the screen, and incongruent trials being sky words presented at the bottom of the screen and ocean words presented at the top of the screen. To test our hypothesis, this base model was compared to a model with an added fixed congruency effect using a likelihood-ratio test. This congruency factor did not explain the data significantly better than the base model did, neither for reaction times, $\chi^2(1)=0.003$, $p=.96$, $\beta=63.84$, $t=0.07$, CI = [−1653.36, 1781.04], nor for the accuracy rates, $\chi^2(1)=0.45$, $p=.50$, $\beta=0.013$, $t=0.66$, CI = [−0.026, 0.052]. Mean solving times and percentages of correctly solved anagrams for each condition can be found in Table 1.

In this experiment, we did not find evidence for the idea that activating location-specific experiential traces would facilitate lexical access to words presumably associated with the corresponding traces. One reason may be that the shift in spatial attention triggered by presenting a stimulus at the top or the bottom of a normal computer screen was simply not strong enough to yield the expected congruency effects. These findings fit with the results from Pecher et al. (2010), showing that merely placing words at the top or the bottom of the screen does not influence word processing sufficiently. In the next experiment we therefore used a background picture that would emphasize the vertical dimension.

Experiment 2

Method

Participants. In this experiment, 28 German native students from the University of Tübingen, with an age ranging from 19 to 34 years ($M=24.14$, $SD=3.70$), participated for money or course credit. Twenty-one of them were female, and none of them had participated in the first experiment. All participants gave written informed consent.

Materials. We used the anagrams from Experiment 1. However, instead of a blank screen, a pictogram of a person was shown in the middle of the screen. This pictogram was present during the presentation of the fixation cross, as well as during anagram presentation itself (see Figure 1). The pictogram disappeared when participants entered their solution, but reappeared when the next fixation cross was shown. By presenting the anagrams either above or below the pictogram of a person, we tried to emphasize the vertical dimension of the computer screen.

Procedure and design. Procedure and design were the same as those in Experiment 1.

Results and discussion

The data were analysed in the same way as in Experiment 1. Outlier elimination reduced the data set by less than 3%. As in Experiment 1, a model with congruency did not

Table 2. Mean solving times and percentage of correctly solved anagrams in Experiment 2.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	8462 (937)	80 (2.86)	7990 (874)	81.12 (3.36)
Sky words	7943 (937)	80 (2.86)	7189 (874)	80 (3.36)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).

Table 3. Mean solving times and percentage of correctly solved anagrams in Experiment 3.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	8666 (1175)	83.81 (2.66)	6463 (1149)	81.67 (3.25)
Sky words	7344 (1175)	84.05 (2.66)	8149 (1149)	81.67 (3.25)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).

predict the data significantly better than a model without it: For reaction times, $\chi^2(1)=0.02, p=.87, \beta=-118.2, t=-0.15$, CI=[−1719.15, 1482.69]; for accuracy rates, $\chi^2(1)=0.14$, $p=.70$, $\beta=-0.006$, $t=-0.38$, CI=[−0.038, 0.026]. Mean solving times and percentage of correctly solved anagrams in Experiment 2 are displayed in Table 2.

Our attempt to emphasize the vertical dimension did not yield any significant congruency effects in this experiment. Either the vertical dimension was still not emphasized enough, or the particular background picture used in this experiment was not ideal. By being a picture of a person it may have primed words that are not related to the sky or ocean category, making solving the anagrams hard even in the compatible conditions. In addition, the picture may have prevented participants from "simulating" the referents of the solution words at the presentation location because there was no match of location relative to the background picture (a fish is not usually encountered below a person). In the next experiment, we therefore presented a background picture that fits better the theme of our solution words.

Experiment 3

Method

Participants. Twenty-eight German native students from the University of Tübingen, with an age ranging from 18 to 44 years ($M=25.07$, $SD=6.51$) participated in the third experiment for money or course credit. Twenty-two of them were female, and none of them had participated in the other two experiments. All participants gave written informed consent.

Materials. We used the anagrams from Experiment 1. However, instead of a blank screen (Experiment 1) or the pictogram of a person (Experiment 2), a drawing of a sailboat was shown in the middle of the screen during presentation of the fixation cross as well as during anagram presentation itself (see Figure 1). As in Experiment 2, we wanted to emphasize the vertical dimension, but this time with a picture that is a better fit with the ocean/sky theme of our solution words.

Procedure and design. The procedure and design were the same as those in Experiment 1.

Results and discussion

Statistical analysis was the same as those in Experiment 1. Outlier elimination reduced the data set by 3.09%. This time, the model with the congruency effect showed a better fit than the base model, $\chi^2(1)=4.19$, $p<.05$, $\beta=1677$, $t=2.07$, CI=[86.37, 3267.71], regarding reaction times. This was not the case for accuracy rates, $\chi^2(1)=0.007$, $p=.94$, $\beta=-0.001$, $t=-0.08$, CI=[−0.034, 0.031]. Mean solving times and percentage of correctly solved anagrams for Experiment 3 are displayed in Table 3. It seems that providing a background picture that matches the sky and ocean theme led to facilitation in compatible trials and/or interference in incompatible trials. This finding suggests that activating experiential traces indeed may activate related concepts and thus facilitate anagram solution processes, as we had predicted. However, since we did not find a significant influence of congruency in the first two experiments, we decided to try to replicate this finding. In Experiment 4, we therefore repeated the procedure of Experiment 3 but this time we used an even more realistic picture matching the sky and ocean theme—namely, a full screen photography of a horizon splitting the screen in two parts—the lower ocean part and the upper sky part.

Table 4. Mean solving times and percentage of correctly solved anagrams in Experiment 4.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	10,186 (1504)	79.76 (3.39)	7664 (1416)	80.71 (3.52)
Sky words	8378 (1504)	81.19 (3.39)	11,076 (1416)	83.1 (3.52)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).

Experiment 4

Method

Participants. Twenty-eight German native students from the University of Tübingen, with an age ranging from 18 to 42 years ($M=24$, $SD=6.50$) participated in the experiment. Twenty-three of them were female, and none had participated in any of the other experiments. All participants gave informed written consent.

Materials. We again used the same anagrams as those in Experiment 1, but this time with a photograph of a horizon showing the sky in the upper half of the screen and the ocean at the bottom half of the screen. The photograph took up the whole screen (see Figure 1). Anagrams appearing at the top of the screen were therefore shown in the sky, while anagrams appearing at the bottom of the screen were shown on the ocean.

Procedure and design. The procedure and design were the same as those in Experiment 1.

Results and discussion

The analysis was identical to that in the previous experiments. Outlier elimination reduced the data set by less than 3%. As in Experiment 3, the congruency effect improved the fit of the model regarding reaction times, $\chi^2(1)=4.40, p<.05$, $\beta=2519.6$, $t=2.12$, CI=[185.86, 4853.37], but did not yield a better fit for accuracy rates, $\chi^2(1)=0.15, p=.70$, $\beta=0.008$, $t=0.38$, CI=[−0.034, 0.050]. The mean solving times and the percentage of correctly solved anagrams in each condition for the fourth experiment are shown in Table 4.

Providing a realistic background picture of a horizon with the sky above and the ocean below seems to facilitate access to ocean and sky words when the participant's attention is drawn to the lower or upper part of the screen, respectively. Thus, activating experiential traces related to the vertical position (i.e., looking up vs. looking down; focusing attention upwards vs. downwards) indeed activates related concepts and thus facilitates anagram solution times in compatible conditions. However, considering that we did not find significant congruency effects in the first two experiments, it seems that activating experiential traces related to vertical position is not enough. Rather what seems needed are more experiential traces that help reduce the number of concepts associated with the vertical position. In Experiments 3 and 4 these traces were presumably activated by the background picture matching the ocean/sky theme. According to this interpretation, both this experiment and Experiment 3 provide evidence in favour of the idea that activating particular experiential traces activates concepts that are associated with these traces. According to this interpretation, the results are therefore in line with the view that experiential traces may indeed play a functional role for cognitive processing. However, similar to recent studies investigating which actual processes underlie the influence of linguistic stimuli on subsequent perceptual processes (Dudschig & Kaup, 2016; Dudschig, Mackenzie, Strozyk, Kaup, & Leuthold, 2016), we here also need to closer investigate what potential mechanisms underlie the influence of perceptual processes on the processing of linguistic material. Thus, there are a number of alternative explanations of these findings that we need to consider before drawing a final conclusion.

First, according to our explanation of the results, the background picture provided additional experiential traces that together with the traces related to vertical position were strong enough to sufficiently activate the solution concept for facilitating anagram solution times. In principle, however, it also seems possible that the background picture mainly suggested to the participants that the solution words were related to the ocean/sky theme. As a consequence, the solution word was already pre-activated to a certain degree both in compatible and in incompatible conditions. For this reduced set of words, the additional information provided by the location-specific experiential trace may then have sufficed to reach the threshold for lexical access in compatible conditions (Morton, 1969), thus facilitating the anagram-solving task. If so, the background picture did not provide additional experiential traces that helped anagram solution processes (as we assumed in our explanation of the results) but possibly helped by providing more abstract knowledge about the set of potential solution words. To test whether the congruency effects in Experiments 3 and 4 were mainly due to such knowledge concerning a reduced set of solution words we conducted Experiment 5. In this experiment, we instructed participants at the beginning of the experimental session that all solution words in the experiment would fit into either the sky or the ocean category. If abstract knowledge about the theme of the solution words drove the effects in Experiments 3 and 4 then we should again find a congruency effect, even when no background picture is being presented.

The second alternative explanation questions the need of activating experiential traces related to the vertical position (looking up vs. down; focusing attention upwards vs. downwards). In principle it seems possible that the pictures helped participants by supporting experiential simulations of the target entities in compatible conditions. For instance, when the anagram was presented below the sailboat or on the ocean part of the screen, participants may have simulated entities typically encountered at this location (e.g., a fish, a shark, or a diver) with the result being that the corresponding concepts get activated, and solution of the anagram-solving task is being facilitated. According to this explanation, the experiential traces provided by the picture are already sufficient to facilitate anagram solution times. In such a case, the experiential traces related to vertical position are not necessary. Congruency effects come about because participants' attention is drawn to the compatible part of the background picture in compatible trials (i.e., for a sky word to the sky part of the picture, for an ocean word to the ocean part of the picture). This alternative explanation seems even more relevant, as the proposed facilitation mechanisms may even be fully amodal in the sense that the compatible part of the background picture (sky vs. ocean) may simply have primed concepts associated with the sky or ocean and thus facilitated anagram solution times. To find out more about the viability of this alternative explanation, we conducted Experiment 6. In this experiment, we presented participants with the anagrams in the centre of the screen. In compatible trials we presented the compatible part of the picture full screen as background (i.e., the sky part of the picture for a sky word; the ocean part of the picture for an ocean word), and in incompatible trials we presented the anagrams on incompatible background pictures. If our alternative explanation is correct, then we should again find congruency effects even though the anagrams are presented in the centre of the screen in both compatible and incompatible trials.

Experiment 5

This experiment was basically the same as Experiment 1, with the same anagrams appearing either at the top or at the bottom of the screen and no background picture. However, in the instructions at the beginning of the experiment, we told participants that all solution words were related to either the ocean or the sky, to narrow down the set of possible solution words. If this abstract knowledge concerning the theme of the solution words is enough to lead to congruency effects with presentation location, then we should find congruency effects in this experiment. In contrast, if the congruency effects observed in Experiments 3 and 4 were indeed due to the background picture providing additional relevant cues for anagram solution, then we should not find a congruency effect in this experiment, in which there was no background picture.

Method

Participants. Twenty-eight German native students from the University of Tübingen, with an age ranging from 19 to 31 years ($M=23.32$, $SD=3.09$) participated in the experiment. Twenty-two of them were female, and none had participated in any of the other experiments. All participants gave informed written consent.

Materials. The materials were the same as those in Experiment 1, except that we informed participants in the instruction that all solution words belonged to the ocean/sky theme.

Procedure and design. The procedure and design were the same as those in Experiment 1.

Results and discussion

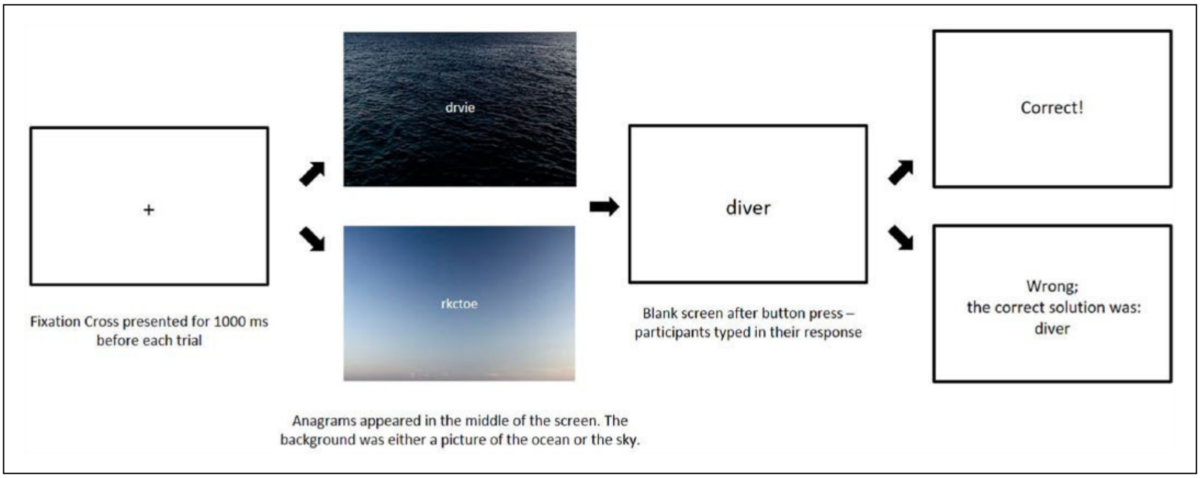
We used the same analysis as that in previous experiments. Outlier elimination reduced the data set by less than 3%. Unlike in Experiments 3 and 4, the model with the congruency effect was not a significantly better fit to the reaction times data than the base model was, $\chi^2(1)=0.03$, $p=.86$, $\beta=115.1$, $t=0.19$, CI = [−1092.75, 1322.92], in this experiment. There was no congruency effect for the accuracy rates, $\chi^2(1)=1.14$, $p=.29$, $\beta=0.021$, $t=1.06$, CI = [−0.017, 0.059]. The mean solving times and the percentage of correctly solved anagrams in each condition for the fifth experiment are shown in Table 5.

Narrowing down the set of potential solution words to only words that fit either the sky or the ocean category did not prove to be the main reason why we found a significant congruency effect in Experiments 3 and 4. Although mean solving times were lower in this experiment than in the others, suggesting that our manipulation worked, we did not find a significant congruency effect without a background picture. This allows ruling out the first of our two alternative explanations for the results observed in Experiments 1 through 4. Thus the question arises whether the results observed so far stem from some kind of (semantic) priming elicited by the part of the background picture on which the anagrams were being presented: Anagrams appearing above the boat (Experiment 3) or in the sky (Experiment 4) are solved faster because the part of the picture they appear on (semantically) primes words like "*sun*", "*cloud*", or "*helicopter*", while anagrams appearing below the boat (Experiment 3) or in the ocean (Experiment 4) prime words like "*dolphin*" and "*diver*" or "*tuna*". This would provide an explanation for our results without the need to assume the activation of experiential traces. To test this explanation, we used the upper and

Table 5. Mean solving times and percentage of correctly solved anagrams in Experiment 5.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	6481 (577)	80.48 (3.12)	6608 (753)	79.05 (2.78)
Sky words	5816 (577)	80.48 (3.12)	6239 (753)	83.10 (2.78)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).



Fixation Cross presented for 1000 ms before each trial

Anagrams appeared in the middle of the screen. The background was either a picture of the ocean or the sky.

Blank screen after button press – participants typed in their response

Figure 2. The structure of a single trial in Experiment 6. After the feedback slide, the next trial started. To view this figure in colour, please visit the online version of this Journal.

lower parts of the background picture from Experiment 4 as separate background images, this time showing the anagrams in the centre of the screen. If the priming explanation is correct, and experiential traces related to vertical position are not needed for finding a congruency effect, then we should find a congruency effect in the next experiment, even though the anagrams were presented in the centre of the screen in all trials. In contrast, if our original explanation of the results is correct, and experiential traces related to the vertical position are relevant for the observed congruency effects, then no congruency effect should be observed.

Experiment 6

Method

Participants. Twenty-eight German native students from the University of Tübingen, with an age ranging from 19 to 35 years (M=23.61, SD=4.44), participated in the experiment. Twenty-three of them were female, and none had participated in any of the other experiments. All participants gave informed written consent.

Materials. The same anagrams were used as those in previous experiments. The photograph of a horizon that was used in Experiment 4 was divided into two separate pictures, one showing only the sky, and the other showing only the ocean (see Figure 2). One of these pictures was shown during presentation of the fixation cross and presentation of the anagrams and took up the whole screen. Instead of manipulating the presentation location of the anagrams, the sky picture was shown whenever an anagram would have appeared at the top of the screen, and the ocean picture was shown whenever an anagram would have appeared at the bottom of the screen in previous experiments. Thus, all anagrams appeared in the middle of the screen, while our manipulation of presentation location was defined by which picture was shown as a background. In order to make sure that the picture halves could indeed be identified by the participants as ocean and sky, respectively, we printed out the pictures and showed them to 15 people who did not participate in any of the six experiments reported in this manuscript. We asked these participants to name what they saw on the pictures. All 15 participants correctly identified the ocean and the sky picture, indicated by their responses “ocean” or “water” to the ocean picture and “sky” to the sky picture.

Procedure and design. The procedure and design were the same as those in Experiment 1, except that instead of manipulating presentation location, we manipulated the presented background picture (sky vs. ocean).

Table 6. Mean solving times and percentage of correctly solved anagrams in Experiment 6.

Word category	Presented at the top		Presented at the bottom	
	Mean solving times (ms) (CI)	Percentage correct (CI)	Mean solving times (ms) (CI)	Percentage correct (CI)
Ocean words	7038 (881)	73.57 (3.22)	7729 (868)	79.19 (3.31)
Sky words	6973 (881)	74.05 (3.22)	7051 (868)	75.95 (3.31)

Note: Confidence intervals were calculated according to Masson and Loftus (2003).

Results and discussion

The statistical analysis was the same as that in previous experiments. Outlier elimination reduced the data set by less than 3%. The model with congruency did not predict the solving times significantly better than the model containing only the main effect for anagram length, $\chi^2(1)=0.18$, $p=.68$, $\beta=-255.4$, $t=-0.42$, CI=[−1452.33, 941.47]. The same was true for accuracy rates, $\chi^2(1)=0.04$, $p=.85$, $\beta=-0.004$, $t=-0.19$, CI=[−0.043, 0.035]. The mean solving times and the percentage of correctly solved anagrams in each condition for the sixth experiment are shown in Table 6.

Since we did not find an effect of congruency in this experiment, our second alternative explanation alluding to semantic priming can be ruled out. In fact, we can even rule out the possibility that the background picture alone provided sufficient experiential traces for facilitating access to the target concepts. Rather it seems that experiential traces related to the vertical position are crucial for finding location-specific congruency effects with an anagram-solving task.

To obtain more information concerning the effects observed in the six reported experiments, we conducted additional analyses involving all or a subpart of the experiments as well as some analyses involving measures of language statistics.

Additional post hoc analyses across experiments

Influence of experiment

To test whether the congruency effect was influenced by the experiment, we pooled the data of all experiments and performed a likelihood-ratio test comparing a base model with word length and congruency as fixed effects and the same random effects structure as that previously described to a model with an additional main effect of experiment, which provided a better fit, $\chi^2(5)=58.02$, $p<.001$. When we added an interaction between congruency and experiment, the fit did not improve further, $\chi^2(5)=9.26$, $p=.10$. So the experiment did have an influence; however, there was no significant interaction between congruency and experiment across all experiments.

Pooling together the data from Experiments 3 and 4, containing both a fitting background picture and spatial information, and comparing them to all other pooled experiments, of which neither contained both of these cues, the model with the interaction between experiment and congruency proved to be the best fit for the data, $\chi^2(3)=7.87$, $p<.01$. This supports the idea that only the presence of both a matching background picture and a manipulated presentation location resulted in significant location-specific congruency effects.

As a manipulation check for Experiment 5 we also tested it against all other experiments. As expected, the model with a main effect for experiment (Experiment 5 vs. all other experiments) did fit the data best, $\chi^2(1)=22.65$, $p<.001$, supporting the idea that narrowing down potential solution words to the sky/ocean theme did in fact reduce the solving times for anagrams, while not modulating the congruency effect, $\chi^2(1)=0.55$, $p=.46$.

Influence of language statistics

As already mentioned in the introduction, some embodiment effects have been shown to be explainable solely by language association statistics. To test for such a possibility in our experimental set-up, we conducted an analysis similar to the one used in Goodhew et al. (2014): We used a German corpus consisting of roughly 880 million words (sDeWaC, see Faaß & Eckart, 2013) and used the R-package *LSAfun* (Günther, Dudschig, & Kaup, 2015) to calculate how often our solution words appeared together with the words "up" and "above" on the one hand and "down" and "under" on the other hand within a window of five words around the target words. As a frequency measure we added the log frequencies for occurring together with "up" and "above", as well as those for occurring together with "down" and "under". We then calculated a difference score by subtracting the latter from the former. A positive log frequency difference thus indicates that the target word more often appears with words referring to the upper vertical space whereas a negative log frequency differences means that the target word appeared more often with words referring to the lower vertical space (for exact values see the Appendix). We then tested whether our word categories (ocean vs. sky words) differed regarding language collocation. Using a student's t-test for unpaired samples we did indeed find a significant difference between our word categories, $t(45.29)=3.29$, $p<.01$, with words from the sky category being associated more strongly with the words “*up*” und “*above*” ($M_{log\_diff}=0.96$) than the ocean words ($M_{log\_diff}=0.44$). Since the log frequency difference in both groups was a positive, both categories were associated more strongly with the words “*up*” and “*above*” than with the words “*down*” and “*under*” but this tendency was significantly stronger for sky than for ocean words. Thus, language statistics (in particular linguistic co-occurrences) could in principle also account for our observed effects of vertical position. In a second step, we then investigated whether the observed differences in linguistic co-occurrence indeed are able to predict the results observed in our experiments: As a base model we started with a main effect for word length, random intercepts for subjects and items, and by-item and by-subject random slopes for presentation location. Adding presentation location and the log frequency difference as main effects did not improve the fit—$\chi^2(2)=0.07$, $p=.97$, $\beta_{log\_diff}=274.62$, $t_{log\_diff}=0.23$, $CI_{log\_diff}=[-2045.89, 2595.14]$, $\beta_{location}=38.69$, $t_{location}=0.11$, $CI_{location}=[-635.81, 713.19]$—and the same was true for an added interaction effect between log frequency difference and presentation location—$\chi^2(3)=3.45$, $p=.33$, $\beta=-969.5$, $t=-1.83$, $CI=[-2008.57, 69.56]$. As we only found a location-specific congruency effect in Experiments 3 and 4, we repeated the same analysis for both of these experiments, with similar results: Neither the addition of the main effects—$\chi^2(2)=0.40$, $p=.82$, $\beta_{log\_diff}=-214.6$, $t_{log\_diff}=-0.16$, $CI_{log\_diff}=[-2777.81, 2348.63]$, $\beta_{location}=368.9$, $t_{location}=0.60$, $CI_{location}=[-840.22, 1577.99]$—nor the interaction in addition to the main effects—$\chi^2(3)=5.80$, $p=.12$, $\beta=-2166.8$, $t=-2.33$, $CI=[-3991.82, -341.76]$—improved the fit of the model. These results therefore suggest that linguistic co-occurrence indeed differs for the two categories at hand but that these differences cannot explain the full pattern of the results observed in the present study.

General discussion

We conducted a series of six experiments in which participants solved anagrams with solution words belonging to the ocean or sky theme. In contrast to the manifold studies investigating the influence of language processing on subsequent spatial processing in perceptual or motor tasks, the current study investigated the reversed direction of influence: Does the activation of spatial experiential traces facilitate linguistic processing of words referring to entities that are typically located in a compatible spatial position in the real world? To test the influence of spatial experiential traces we manipulated the presentation location of the anagrams. In half of the trials, the anagrams were presented at the top of the screen, in the other half at the bottom of the screen. Based on the hypothesis that activating location-specific experiential traces should activate the concepts that are associated with these traces, we predicted to find location-specific congruency effects, with faster anagram solution times when the anagram’s position on the screen matched the typical position of the respective referent in the world. We found significant congruency effects, but only in Experiments 3 and 4, in which the anagrams were presented with a background picture that was associated with the ocean/sky theme. In Experiments 1 and 2, in which no or an unrelated background picture was presented, we did not observe significant congruency effects. These findings are in line with previous studies suggesting that merely activating spatial traces does not result in facilitated linguistic processes (e.g., Pecher et al., 2010). However, of course the question arises in which way the background pictures in Experiments 3 and 4 provided the necessary context for finding the predicted congruency effects. The most likely explanation of the observed results in our view is that a rich set of experiential traces is necessary to pre-activate the target concepts sufficiently to find facilitation in an anagram-solving task. Location information seems to be a particularly relevant cue in this task. The results of two control experiments allowed ruling out two alternative explanations. Neither abstract knowledge about the theme of the solution words presented at the beginning of the experiment (Experiment 5) nor a set of non-spatial experiential traces (Experiment 6) sufficed to provide the necessary context for the predicted congruency effect.

Thus, our results suggest that spatial experiential traces can activate associated concepts, but only within a supporting context, in our case a background picture fitting the theme of the solution words that enhances these spatial traces. Neither spatial traces alone nor a fitting background picture alone systematically influence anagram-solving processes. We therefore interpret the observed congruency effects as reflecting a mixture of automatic activation processes and the integration of the context information provided by the background picture. More specifically, drawing participants’ attention to particular locations on the screen activates location-specific experiential traces, which in turn activate concepts that are associated with these locations. However, by themselves these location-specific experiential traces are simply not powerful enough to activate potential solution words to such a strong degree that this would facilitate solution times in a particular anagram task. We therefore did not find congruency effects in our first experiment. Providing a background picture that does not fit with the solution words thematically, as in Experiment 2, also did not lead to reduced solution times in compatible trials, because the activated location information together with the given context of a person may have pre-activated other concepts like “*hat*” or “*shoe*”, but not our solution words. Matching context pictures alone without the activation of locational traces did not facilitate the solution of words that matched the presented picture (sky or ocean) for similar reasons (Experiment 6). Finally, providing abstract knowledge to the participants about the set of possible solution words at the beginning of the experimental session did not have the same effect as a matching background picture during anagram solution, suggesting that materials activating relevant non-linguistic experiential traces are particularly effective as additional cues to the target concepts in anagram-solving tasks.

We interpreted the differences in results obtained in Experiments 3 and 4 on the one hand and Experiment 5 on the other hand as reflecting the difference between providing abstract linguistic knowledge (Experiment 5) versus non-linguistic experiential traces (Experiments 3 and 4) as further cues towards the theme of the target words. However, in principle it also seems possible that temporal characteristics of the experimental procedure are responsible for the differences. In Experiments 3 and 4, the additional cues (visual experiential traces) were presented repeatedly and simultaneously to the anagram task, whereas in Experiment 5, the additional information (instruction about theme of solution words) was presented only once at the beginning of the experiment. Thus, possibly, abstract linguistic knowledge about the theme of the solution words would suffice to find congruency effects in case this information were given in each trial and simultaneously to the anagram task. We cannot rule out this hypothesis on the basis of the data presented here. To test this, one would need to present participants with the theme of the solution words in each trial, for instance by presenting the anagrams in the context of an instruction to find a solution word that fits the ocean/sky theme.

Our interpretation of the results is fully in line with grounded models of cognition in that it exclusively refers to modal meaning representations. In recent years many authors have suggested hybrid forms of meaning representations to underlie most cognitive processes (see e.g., Dove, 2009). According to hybrid accounts, modal meaning representations in the form of re-activated experiential traces are only one kind of meaning representation that people have available, which is only used under certain conditions, for instance in contexts that strongly suggest to the participants to engage in experiential simulations of objects and events. Applied to our results, hybrid accounts could propose that providing a background picture may have prompted our participants to engage in some kind of visual simulation process, which together with the positional information helped to solve the anagrams in compatible trials or hindered solution in incompatible trials. This would explain why significant congruency effects were observed in Experiments 3 and 4 but not in Experiments 1 and 5, where no background picture was being presented, and participants therefore presumably processed the anagrams in a strictly amodal way. However, problematic for such an account is the fact that no congruency effects were observed in Experiments 2 and 6 although background pictures were present during anagram solution. Nevertheless future studies should further investigate the viability of hybrid accounts for explaining anagram solution times. One possibility would be to present participants with multimodal stimuli in an intermediary task between anagram trials to trigger their willingness to engage in experiential simulation processes.

In our experiments, spatial information related to a referent's typical position in vertical space played a prominent role in investigating the influence of experiential traces on anagram solution times. As an extension to our results it would be interesting to look at other experiential dimensions. For instance, it has been shown that facial expressions are strongly linked to emotions (e.g., see Adelmann & Zajonc, 1989). One could manipulate the participants' facial expression, for instance by letting them put a pen between their lips (similar to frowning; negative facial expression) or between their teeth (similar to smiling; positive facial expression; see Strack, Martin, & Stepper, 1988; but see: Wagenmakers et al., 2016), and then let them solve anagrams of positive or negative emotion words. If solution times were faster in compatible conditions, this would provide further evidence that experiential information can activate certain concepts. Similar approaches are conceivable using other non-visual senses, for example sounds or smells, during anagram-solving tasks and investigating their influence on solution times.

In addition to providing insight regarding how word meaning and concepts are represented, the methodology used here could also help to get a better understanding of complex cognitive processes, as investigated in the problem solving literature. Influences of different aspects, like single letter cues (Witte & Freund, 2001) or phonemic cues (Fink & Weisberg, 1981), on anagram solving have already been shown. In this study we could also show that cues from other modalities linked to language, in this case a matching picture and the position of the anagram on the screen, can facilitate anagram solving. The underlying mechanisms and the time course of this influence remain to be investigated in future studies. It is still unclear whether this facilitation can be generalized to other word categories, whether it is stronger for easier or more difficult anagrams, and how it interacts with different anagram-solving strategies. Novick and Sherman (2003, 2008), for example, showed that skilled anagram solvers apply different strategies from poor anagram solvers, and that expertise may influence the helpfulness of different cues or distractors. It would be interesting to investigate whether experiential traces, like those used in the experiments reported here, would have a different impact for different expertise groups.

Conclusion

In a series of six experiments we found evidence that anagram solving can be influenced by the interplay of position on the screen and background picture, suggesting an influence of sensorimotor processes on concept activation. In this way our study differs from previous studies that mostly focused on an influence of processing certain words or sentences on subsequent sensorimotor processes. Our results allowed ruling out that the observed effects are mainly due to language collocation or semantic priming, thus providing further evidence for a multimodal representation of concepts. We suggest that anagram-solving tasks will prove useful for research on linguistic and conceptual processing in future research, because they provide a simple measure of accessibility of certain words or concepts in a productive rather than a receptive task. In addition, the results of our study are relevant not only to research on linguistic and conceptual processing but also for research on problem solving, showing a clear influence of multimodal information on an anagram-solving task.

Acknowledgements

We thank Fritz Günther for calculating the word co-occurrences using the sDeWaC corpus.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the German Research Foundation (DFG) as part of the Collaborative Research Center "SFB 833"; Project B4 of Barbara Kaup and Hartmut Leuthold.

References

Adelmann, P. K., & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, *40*(1), 249–280.

Areshenkoff, C. N., Bub, D. N., & Masson, M. E. (2017). Task-dependent motor representations evoked by spatial words: Implications for embodied accounts of word meaning. *Journal of Memory and Language*, *92*, 158–169.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Brookshire, G., Casasanto, D., & Ivry, R. (2010). Modulation of motor-meaning congruity effects for valenced words. In *32nd Annual Meeting of the Cognitive Science Society (CogSci 2010)* (pp. 1940–1945). Cognitive Science Society.

Bub, D. N., Masson, M. E., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*(1), 27–58.

Dominowski, R. L., & Ekstrand, B. R. (1967). Direct and associative priming in anagram solving. *Journal of Experimental Psychology*, *74*(1), 84–86.

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*(3), 412–431.

Dudschig, C., & Kaup, B. (2016). Is it all task-specific? The role of binary responses, verbal mediation, and saliency for eliciting language-space associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication, http://dx.doi.org/10.1037/xlm0000297

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). Do task-irrelevant direction-associated motion verbs affect action planning? Evidence from a Stroop paradigm. *Memory & Cognition*, *40*(7), 1081–1094.

Dudschig, C., Mackenzie, I. G., Strozyk, J., Kaup, B., & Leuthold, H. (2016). The sounds of sentences: Differentiating the influence of physical sound, sound imagery, and linguistically implied sounds on physical sound processing. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(5), 940–961.

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "sun" and looking up: The influence of language on saccadic eye movements in the vertical dimension. *PloS one*, *8*(2), e56872.

Dudschig, C., de la Vega, I., & Kaup, B. (2015). What's up? Emotion-specific activation of vertical space during language processing. *Acta Psychologica*, *156*, 143–155.

Ellis, J. J., Glaholt, M. G., & Reingold, E. M. (2011). Eye movements reveal solution knowledge prior to insight. *Consciousness and Cognition*, *20*(3), 768–776.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down object words orient attention to the objects' typical location. *Psychological Science*, *19*(2), 93–97.

Faaß, G., & Eckart, K. (2013). SdeWaC–a corpus of parsable sentences from the web. In *Language processing and knowledge in the Web* (pp. 61–68). Berlin, Heidelberg: Springer.

Fink, T. E., & Weisberg, R. W. (1981). The use of phonemic information to solve anagrams. *Memory & Cognition*, *9*(4), 402–410.

Fodor, J. A. (1975). *The language of thought* (*Vol. 5*). Cambridge, MA: Harvard University Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Goodhew, S. C., McGaw, B., & Kidd, E. (2014). Why is the sunny side always up? Explaining the spatial mapping of concepts by language use. *Psychonomic Bulletin & Review*, *21*(5), 1287–1293.

Günther, F., Dudschig, C., & Kaup, B. (2015). LSAfun-An R package for computations based on Latent Semantic Analysis. *Behavior Research Methods*, *47*(4), 930–944.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25*(1), 236–242.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, *38*(7), 1033–1050.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: Automatic activation of location information during word processing. *Psychonomic Bulletin & Review*, *18*(6), 1180–1188.

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science*, *39*(8), 1764–1801.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*(1), 96–104.

Masson, M. E., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *57*(3), 203–220.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up associations between affect and vertical position. *Psychological Science*, *15*(4), 243–247.

Morton, J. (1969). Interaction of information in word recognition. *Psychological review*, *76*(2), 165–178.

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(2), 351–382.

Novick, L. R., & Sherman, S. J. (2008). The effects of superficial and structural information on online problem solving for good versus poor anagram solvers. *The Quarterly Journal of Experimental Psychology*, *61*(7), 1098–1120.

Pecher, D., Van Dantzig, S., Boot, I., Zanolie, K., & Huber, D. E. (2010). Congruency between word position and meaning is caused by task-induced spatial attention. *Frontiers in Psychology*, *1*, 30. doi:10.3389/fpsyg.2010.00030

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*(3), 793–797.

Schuil, K. D., Smits, M., & Zwaan, R. A. (2013). Sentential context modulates the involvement of the motor cortex in action language processing: An FMRI study. *Frontiers in Human Neuroscience*, *7*, 1–13.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*(2), 297–312.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*(5), 768–777.

Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B.Jr.,… Zwaan, R. A. (2016). Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, *11*, 917–928.

Witte, K. L., & Freund, J. S. (2001). Single-letter retrieval cues for anagram solution. *The Journal of General Psychology*, *128*(3), 315–328.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. Zwaan (Ed.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 224–245). New York: Cambridge University Press.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.

Appendix

Table A1. Anagrams, solution words, translation, number of letters, frequency class, and log frequency difference (see Goodhew et al., 2014) of the "ocean words" used in the experiment.

Anagram	Solution word (in German)	English translation	Number of letters	Frequency class	Log frequency difference
Aegl	Alge	algae	4	16	0.47000363
Tgna	Tang	kelp	4	15	0.49247649
Sdna	Sand	sand	4	10	0.16945505
Arnke	Anker	float	5	13	0.44468582
Oetrt	Otter	otter	5	16	0.78845736
Lshca	Lachs	salmon	5	13	0.72054615
Rbebo	Robbe	seal	5	14	−0.2006707
Kebsr	Krebs	crab	5	11	0.45534647
Wcrka	Wrack	ship wreck	5	13	0.08701138
Hrgine	Hering	herring	6	13	1.09861229
Qaulel	Qualle	jellyfish	6	17	0.40546511
Hmurem	Hummer	lobster	6	13	−0.0645385
Asuetr	Auster	oyster	6	16	1.01160091
Dnefil	Delfin	dolphin	6	15	1.09861229
Kenark	Kraken	kraken	6	16	0.13353139
Graleen	Garnele	shrimp	7	18	0.91629073
Mlesuch	Muschel	clam	7	15	0.49643689
Kroelal	Koralle	coral	7	17	0.22314355
Tacehur	Taucher	diver	7	13	0
Opktous	Oktopus	octopus	7	16	1.04982212
Flrloee	*Forelle*	*trout*	7	15	0.81093022
Knarfpe	*Karpfen*	*carp*	7	14	−0.50077529
Potlnnak	Plankton	plankton	8	16	0.69314718
Kbaljeua	Kabeljau	cod	8	15	0.98082925
Hifihsac	Haifisch	shark	8	18	0
Fezcnthsi	Fischnetz	fishing net	9	21	0
Tnsfuchih	Thunfisch	tuna	9	14	0.28768207
Tuacmhsake	Tauchmaske	diving mask	10	21	0
Srhocnlhce	Schnorchel	snorkel	10	17	0.69314718
Thistncinef	Tintenfisch	squid	11	15	0.31845373

Note: Words in italics were not taken from Pecher et al. (2010).

Table A2. Anagrams, solution words, translation, and number of letters of the “sky words” used in the experiment.

Anagram	Solution word	English translation	Number of letters	Frequency class	Log frequency difference
Mdon	Mond	moon	4	11	1.07149905
Eleu	Eule	owl	4	14	0.76546784
Reab	Rabe	raven	4	13	0.42744401
Snoen	Sonne	sun	5	9	−0.04231943
Wekol	Wolke	cloud	5	12	1.38295546
Koetm	Komet	comet	5	15	0.82098055
Alred	Adler	eagle	5	11	1.24764787
Snret	Stern	star	5	10	−0.42034317
Giere	Geier	vulture	5	14	1.33500107
Blalno	Ballon	balloon	6	13	1.12458778
Mtoeer	Meteor	meteor	6	16	1.09861229
Rkatee	Rakete	rocket	6	12	0.57536414
Pnetla	Planet	planet	6	12	0.16475523
Shptce	Specht	woodpecker	6	14	0.84729786
Sotrhc	Storch	stork	6	14	0.81831032
Hachbit	Habicht	hawk	7	17	0.77318989
Krloibi	Kolibri	hummingbird	7	17	0.22314355
Dharecn	Drachen	kite	7	13	0.87457228
Tranodo	Tornado	tornado	7	14	1.75785792
Busrdsa	Bussard	buzzard	7	17	1.54044504
Wlaetll	Weltall	universe	7	13	1.42711636
Zinpepel	Zeppelin	zeppelin	8	14	1.70474809
Flzguueg	Flugzeug	airplane	8	10	1.52493123
Sialltte	Satellit	satellite	8	13	3.64683122
Asanrottu	Astronaut	astronaut	9	14	0.11778304
Frkwueeer	Feuerwerk	fireworks	9	11	1.51059208
Hpoiketrle	Helikopter	helicopter	10	12	1.64222774
Reogeegnnb	*Regenbogen*	*rainbow*	10	14	0.53062825
Rktoehcelhn	Rotkehlchen	robin	11	16	0.13353139

Note: Words in italics were not taken from Pecher et al. (2010).

B. Study 2:

Copyright Notice

Reprinted from *Psychological Research,* Berndt, E., Dudschig, C. & Kaup, B. (2018). Green as a cbemcuru: modal as well as amodal color cues can help to solve anagrams. Copyright 2018, with permission from Springer Link to use as part of this dissertation. The official citation that should be used in reference to this article is:

Berndt, E., Dudschig, C., & Kaup, B. (2018). Green as a cbemcuru: modal as well as amodal color cues can help to solve anagrams. *Psychological Research*. https://doi.org/10.1007/s00426-018-1055-y

Psychological Research
https://doi.org/10.1007/s00426-018-1055-y

ORIGINAL ARTICLE

Green as a cbemcuru: modal as well as amodal color cues can help to solve anagrams

Eduard Berndt[1] · **Carolin Dudschig**[1] · **Barbara Kaup**[1]

Received: 28 March 2018 / Accepted: 7 July 2018

Abstract

Embodied cognition theories have been getting much support in recent years from studies showing that multimodal experiential traces are activated during language comprehension. However, there are almost no studies examining this influence in the opposite direction. Here, we investigated the influence of modal (physical color patch) and amodal (color word) cues on anagram solving times. We manipulated the association between the color cue and the solution word's referent color (e.g., finding the solution word "cucumber" for the anagram "cmrbucue" should be facilitated by the word "green" or a green color patch). In a third experiment, both cues were combined: a color word was presented inside a color patch before the anagram appeared. We indeed observed priming effects: anagrams were solved faster when the preceding color patch or color word matched the solution word's referent compared to a mismatching color patch or color word. When combining these cues, a priming effect only was found when both color word and color patch matched the solution word's referent. These results further strengthen the notion that multimodal experiential traces play an important role in language comprehension and expand upon the results of earlier studies on anagram solution tasks.

Introduction

Theories of embodied cognition have been getting much attention in recent years. In contrast to more traditional propositional theories of language comprehension (e.g., Fodor 1975), that propose that language comprehension is solely based on manipulation of amodal, arbitrary, and abstract symbols, these embodied models assume that language comprehension is multimodal in nature (Barsalou, 1999). According to this view, a word automatically activates a multimodal concept in long-term memory that consists of past experiences the comprehender has made when experiencing or interacting with the word's referent, be they visual, tactile, olfactory, or auditory. This concept then presumably gives rise to a multimodal conceptual representation in working memory that represents the meaning of the linguistic input. This is supposed to hold for single words, but also for words in context (e.g., the sentence or discourse), whereby the context in which single words appear is assumed to alter the activated conceptual working memory representation of the described objects and events.

Zwaan, Stanfield, and Yaxley (2002) have shown that the shape of an object is activated when reading a sentence, and that this shape is sensible to changes in context. Participants were faster to recognize a picture displaying an eagle with outstretched wings after having read a sentence such as "He saw the eagle in the sky" compared to a sentence such as "He saw the eagle in the nest". The reverse is true for a picture of a perched eagle. Similar effects have been observed for other object properties such as orientation (Stanfield & Zwaan, 2001) or associated vertical position (Dudschig, de la Vega & Kaup, 2015; Dudschig & Kaup, 2017; Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013; Lachmair, Dudschig, De Fillipis, de la Vega, & Kaup, 2011). Stanfield and Zwaan (2001), for instance, showed pictures of a pencil in a horizontal or vertical position to their participants after presenting them with sentences such as "the pencil is in the cup" or "the pencil is in the drawer". Picture recognition was affected by the match between the implied orientation in the sentence and the depicted orientation in the picture, with faster response times in the match compared to the mismatch condition. Similarly, Lachmair et al. (2011) used single words that refer to entities that are associated with either the upper or lower vertical space (e.g., "sun" versus

---

✉ Eduard Berndt
eduard.berndt@uni-tuebingen.de

[1] University of Tübingen, Schleichstrasse 4, 72076 Tübingen, Germany

"shoe") (see also Dudschig, de la Vega, De Filippis, & Kaup, 2014). Participants had to respond to these words by an upward or downward-oriented arm movement. Again the results showed that matching motor responses (e.g., responding upwards after encountering the word "sun") were faster than mismatching responses.

There have been mixed results with sentence-picture paradigms concerning another important object property: color. Connell (2007) had participants read sentences such as "John looked at the steak on the plate/in the butchers window", implying the mentioned steak being either brown or red. Afterward, participants saw a picture of a steak that was either brown or red and were asked to decide whether the pictured object was mentioned in the sentence. In contrast to the studies described earlier, response times were actually faster when the picture mismatched the color implied by the sentence. The authors argued that color is not a "primary" object property like form or shape, and can thus be easily ignored, which would be advantageous in the task that was employed. The particularly fast response times in the mismatching condition were explained by assuming that the color in the picture was easier to ignore if it mismatched the color implied by the sentence. However, a direct replication by Zwaan and Pecher (2012) and a conceptual replication with new sentences and pictures by Mannaert, Dijkstra, and Zwaan (2017) yielded faster reaction times after matching sentences compared to mismatching sentences, providing further evidence for multimodal concepts as postulated by Barsalou (1999).

Empirical results such as the ones described in the previous paragraphs are usually taken as evidence for the view that meaning representations are multimodal in nature. The underlying logic is the following: a multimodal conceptual representation should include (modal) information such as shape, orientation, location or color, even though these meaning dimensions are typically not explicitly referred to in the linguistic materials. Propositional representations in contrast, being more abstract and symbolic, do not necessarily represent this kind of implicit modal information (Barsalou, 1999, 2008, but see Kintsch, 1988, 2001). As a consequence, if comprehenders indeed activate such multimodal experiential representations during comprehension, then response times in a subsequent non-linguistic task (e.g., picture recognition, motor responses) should be faster if this task involves features that match rather than mismatch features of the multimodal conceptual representation that was activated during comprehension. Finding such a response time pattern then supports the assumption that comprehenders have indeed activated multimodal conceptual representations of the described objects during comprehension. This in turn is assumed to fit well with the idea that multimodal representations indeed constitute the meaning representations of the respective words (but see Mahon & Caramaazza, 2008).

One important implication that can be derived from theories of embodied cognition has received far less attention in the literature. If the representation of a word's meaning is a multimodal representation (with multimodality originating from experiential traces), then one should observe a priming effect in the reverse direction as well: activation of a sufficient amount of experiential traces should facilitate access to certain words (see Kaup, de la Vega, Strozyk & Dudschig, 2015). This should also help to shed light on the open question whether activation of multimodal traces is merely a by-product of spreading activation upon processing a word, or whether it is an integral part of the conceptual representation. Berndt, Dudschig, and Kaup (2016) tested this assumption using an anagram solving task (see Mills, Boychuk, Chasteen, & Pratt, 2018 for a similar study using a different task). Participants were presented with anagrams of nouns (resulting from scrambling the letters of a word) at the top or at the bottom of the screen. The anagram's solutions were associated with upper or lower vertical space (more precisely words associated with the sky or the ocean, e.g., "cloud" and "shark"). The idea was that presentation location would act as a modal cue activating experiential traces related to the upper or lower vertical space, resulting in activation of associated conceptual representations which in turn would facilitate anagram solving of words related to these concepts. This is a modal type of associative priming where primes are not associated words but modal aspects of the solution word's referent. Associative priming with words as a prime has already been shown for anagram solution tasks (e.g., showing the word "table" as a prime before an anagram of "chair", see Dominowski and Ekstrand, 1967). Berndt et al. (2016) indeed observed faster solution times when the presentation location of the anagram matched the associated position of the solution word's referent, but only when the anagrams were shown on a background picture depicting the ocean on the lower half and the sky on the upper half of the screen. A mere shift of attention toward upper or lower vertical space did not seem to activate enough experiential traces to activate the ocean or sky words.

These results can be considered surprising, given the abstract nature of anagram solving tasks. According to classical problem solving theories (e.g., Newell & Simon, 1972), anagram solving tasks can be classified as well-defined problems. Participants are assumed to create an abstract problem space as a form of internal representation of the current state of the problem (the arrangement of the letters), their possibilities of interacting with the problem (switching the letters around), and the desirable solution of the problem (finding a meaningful noun with the given letters). Given this abstract representation, it seems difficult to imagine how modal cues, as were used in Berndt et al. (2016) would facilitate finding a solution. Kirsh (2009) encouraged a more situated view on problem solving, which could explain these results: he claims that translating a problem into an abstract problem space and back to manipulate our environment when trying to solve a problem would seem ineffective. According to him, we use hints and cues from our environment to engage in every day problem solving. For example, Scribner (1985) observed that milkmen used their delivery case to help them with calculations regarding their orders and that they were even faster using this strategy than students using arithmetic calculations.

In the current study, we aimed at expanding upon the findings of Berndt et al. (2016) using color instead of vertical position as the relevant meaning dimension. In a first experiment, we used color words as amodal cues to find out whether associative priming effects could be extended to the dimension of color. We expected faster solution times for anagrams when the solution word's referent was associated with the color word that was presented before the anagram (showing "green" before showing an anagram for "cucumber").

In a second experiment, instead of color words, we used a single color patch that was shown prior to the anagram. Instead of an amodal cue we thus used a modal cue. If multimodal experiential traces are crucial to conceptual representations, we would expect to find the same pattern as in the first experiment: faster solution times when the presented color matches the color that is associated with the solution word's referent.

In a third experiment, we combined the first two experiments, presenting a color word inside a color patch. The solution word's referent could either be associated with both cues, with none of the cues or with either one of the cues. This way, we wanted to investigate whether conceptual representations were more strongly activated by either amodal or modal cues. If the involvement of sensory systems is only due to cascading activation originating from amodal representations, we would expect a much weaker influence of modal cues compared to amodal cues. If it is the other way round, we would expect modal cues to exert a stronger influence. If both cues are processed similarly, they should cancel each other out and we would only expect to find match effects when both modal and amodal color cues match the color of the solution word's referent.

Experiment 1

In this experiment, we used amodal cues prior to anagram solving. Participants were presented with color words that either matched or mismatched the color that was associated with the anagram's solution word. If amodal cues influence the ease with which an anagram can be solved then we should observe faster solution times in the matching compared to the mismatching conditions.

Participants

Participants were recruited via an E-Mail to students of the University of Tübingen. Participation was voluntary, and participants could enter a lottery to win 1 of 6 20 € vouchers for an online shop. 108 participants completed the experiment (75 female, age ranging from 18 to 46 years, $M = 23.09$).

Material

The experiment was implemented using JsPsych (Version 5.0.3.). It was accessible via a link and could be completed by the participants at home, using a common web browser (see de Leeuw & Motz, 2016 for a validation study of using JsPsych for behavioral experiments in a web browser).

We used 80 German words referring to entities that are typically associated with 1 of 4 colors (green, yellow, red, or brown), with 20 words for each color [e.g., "Gurke (cucumber)" for green, "Tomate (tomato)" for red, "Schokolade (chocolate)" for brown and "Banane (banana)" for yellow]. These word groups did not differ significantly in word length or frequency (both $F\text{s}(3,76) < 1$). Words were presented as anagrams that were created randomly for each participant by scrambling the letters of the word, keeping the first letter in the correct position (e.g., "cmrbucue" could have been an anagram for "cucumber"). Before each anagram, 1 of 4 color words ["Grün (green)", "Gelb (yellow)", "Rot (red)", or "Braun (brown)"] was displayed in black font.

Procedure

The experiment started with 4 short practice trials to familiarize participants with the task, followed by the 80 experimental trials. A centered fixation cross was shown briefly (300 ms) at the start of each trial, followed by a color word, written in black font, for 700 ms that served as an amodal color cue. Then, one of the anagrams was shown in the center of the screen until the participant pressed any key indicating that he or she had solved the anagram. The key press made the anagram disappear from the screen. Participants then typed their solution into a text box that appeared at the anagram's location and confirmed their response with the Enter Key. Afterwards a feedback was shown on the screen, showing the correct solution if participants had not responded correctly (lasting 1200 ms), or a confirmation of success when they did (lasting 1000 ms). Afterwards, the next trial started. The time between the onset of the anagram and the participant's key press was considered to reflect the solving time for the anagram and served as dependent variable in our analyses. The whole experiment took approximately 15–25 min to complete.

Design

We implemented four different lists, with each word being shown with each color word in one of the lists. Every participant saw only one of the lists and thus saw each word only once with one of the color words. Since we had 4 colors, there were 20 congruent (5 per color) and 60 incongruent (15 per color) trials per list. Participants were assigned to one of the lists at random at the start of the experiment, and within a list, each trial was chosen randomly.

Thus, our experiment was a 4 (color word) × 4 (referent color) × 4 (list) design, with color word and referent color as within-participant variables and list as a between-participant variable. For the analyses, we combined the first two variables into a single congruency factor, with congruent trials being the ones where the color word matched the referent color, and incongruent trials being the ones where color word and referent color did not match. As dependent variables, we analyzed anagram solving times (see above) and the percentage of correctly solved anagrams (accuracy).

Results

Before analyzing the data, we excluded participants with less than 66% correct answers (excluding five participants) to make sure that all participants included in the analysis were actually engaged in the task. This seemed particularly important, as the study was a web browser-based study, with participants conducting the task at home. In addition, three of the remaining participants were excluded because the lists received slightly different numbers of participants because of random participant-to-list assignments. To equalize the Ns in the four lists, we excluded the most recent participants from two of the four lists to bring the number of participants down to the number of the lists with the least participants. This resulted in 100 participants (69 female, age ranging from 18 to 39, $M = 22.97$) who were included in the final analysis, 25 per list. Solving times above 20 s were removed as outliers, which reduced our data set by less than 6%. For the analysis of solving times only correct trials were taken into account.

We conducted our analysis with the free statistic software R (Version 3.41). Using the R-package lme4 (Bates, Mächler, Bolker, & Walker, 2015), we tested our hypotheses with linear mixed effects models (LMEMs). Our base model consisted of fixed main effects for anagram length and word frequency, random intercepts for participants and items, as well as by-participant and by-item random slopes for congruency. We then compared the base model to a model that had an additional fixed effect for congruency with a likelihood ratio test. This resulted in a significantly better fit for the model with added congruency for solving times [$\chi^2(1) = 7.34$, $p = .007$, $\beta = -280.62$, $t = -2.77$, CI (−479.50, −81.74)]. The mean solving time for anagrams in matching trials was faster compared to mismatching trials (M_{Match} = 3.184 s, M_{Mismatch} = 3.456 s). The detailed solving times are summarized in Table 1. To make sure that our measure of solving time was adequate and participants did not press a key before they actually had an answer, we conducted the same analysis with the solving time being defined as the time from anagram onset to confirming the response. We again excluded solving times above 20 s (resulting in 6194 trials compared to 6310 trials in the original analysis). This did not change the results [$\chi^2(1) = 11.44$, $p < .001$, $\beta = -309.78$, $t = -3.40$, CI (−488.33, −131.24)]. We did not find any influence of matching versus mismatching prime color on accuracy rates [$\chi^2(1) = 1.78$, $p = .41$, $\beta = 0.005$, $t = 0.58$, CI (−0.012, 0.022)], with mean accuracy rates of 82.33% in matching trials and 81.81% in mismatching trials. Including list as an additional factor did not further improve the fit for the prediction of solving times or accuracy (both ps > 0.6).

Table 1 Mean solving times in ms for each combination of color word and associated color of the solution word's referent and the magnitude of the match effect (mean of solution times of mismatching trials—mean of solution time of matching trials)

Referent's color	Color word				Match effect
	Brown	Yellow	Green	Red	(MM-M)
Brown	3380	3605	3883	3868	405
Yellow	3382	3073	3221	3499	294
Green	3405	3006	3071	3529	242
Red	3359	3384	3330	3211	147

Discussion

The results showed that anagram solution times are faster when the color word cue matched than when it mismatched the solution word's referent (i.e., "green" as a cue before an anagram for "cucumber" lead to faster solution times compared to "red" as a cue). This result can be interpreted as a conceptual replication of the results of Dominowski and Ekstrand (1967). Here, we used color names as cues for words associated with that color instead of semantically associated words priming each other. In the next step, we aimed at investigating whether a priming effect would also occur when instead of color names as amodal linguistic cues a modal, visual cue in form of the color itself is used. Theories of embodied cognition would predict such priming effects, since conceptual representations are assumed to consist of multimodal experiential traces and activating these traces should facilitate access to corresponding concepts. Previous studies have shown associative picture—word priming and word—picture priming effects when participants had to name the target picture or read out loud the target word (Sperber, McCauley, Ragain, & Weil, 1979). However, there are mixed results concerning cross-modal priming effects on anagram solving tasks. Srinivas and Roediger (1990), for example, failed to find priming effects of pictures on a subsequent anagram solving task. Antonietti and Girotti (1991) reported direct priming of written words that were later used as anagrams, but did not find such priming effects when these words were presented auditorily. Rajaram and Roediger (1993) on the other hand did find direct priming effects for auditorily presented words in a similar paradigm. Regarding the influence of color on anagram solving, Mehta and Zhu (2009) found the color red to prime words associated with avoidance motivation and the color blue to prime words associated with approach motivation in an anagram solving task, but Steele (2014) failed to replicate this finding. Thus, on the basis of the available evidence in the literature it is difficult to predict whether modal cues influence anagram solving processes in a similar way as amodal cues do.

Experiment 2

In this experiment, we used modal cues prior to anagram solving. Participants were presented with color patches that either matched or mismatched the color that was associated with the anagram's solution word. If modal cues influence the ease with which an anagram can be solved then we should observe faster solution times in the matching than in the mismatching conditions, as in Experiment 1.

Participants

As in the first experiment, participants were recruited by sending a link via E-Mail to students of the University of Tübingen. We chose a different subgroup of students ensuring that new participants were recruited. Additionally the link was posted in a social media group for psychological experiments. Participation was again voluntary, and participants could enter another lottery to win 1 of 6 20 € vouchers for an online shop. 131 participants completed the experiment (91 female, age ranging from 18 to 55 years, $M = 26.89$).

Material and procedure

The material as well as the procedure was the same as in the first experiment. The only difference being that instead of a color word, participants saw a colored patch before each anagram as modal color prime. It was a colored rectangle sized 400 × 300 pixels that was shown in one of the four colors (green: RGB 0,255,0; red: RGB 255,0,0; brown: RGB 139,69,19; and yellow: RGB 255,255,0) for 700 ms.

Table 2 Mean solving times in ms for each combination of prime color and associated color of the solution word's referent and the magnitude of the match effect (mean of solution times of mismatching trials—mean of solution time of matching trials)

Referent's color	Prime color				Match effect
	Brown	Yellow	Green	Red	(MM-M)
Brown	3614	3920	4142	3775	332
Yellow	3356	3360	3304	3314	– 35
Green	3619	3341	3200	2992	117
Red	3681	3436	3589	3062	507

Design

The design was identical to Experiment 1 except that the factor color word was replaced by the factor prime color, leading to a 4 (prime color) × 4 (referent color) × 4 (list) design, with prime color and referent color as within-participant variables and list as a between-participant variable. The factor's prime color and referent color were combined into a new congruency factor for the analysis, as we did in the first experiment.

Results

We used the same exclusion criteria as in the first experiment: participants with less than 66% correct answers in the experimental trials were excluded (excluding 12 participants). In addition, the *N*s of the 4 lists were equalized, excluding the most recent participants from 3 of the 4 lists to bring the number of participants down to the number of the list with the least participants (excluding 19 participants). This resulted in 100 participants (71 female, age ranging from 18 to 55, $M = 26.6$) who were included in the final analysis, 25 per list. Removing solving times above 20 s reduced our dataset by less than 6%. For the analysis of solving times only correct trials were taken into account.

We used the same statistical procedure as in the first experiment, comparing a linear mixed effects model without the congruency factor to the same model with an added congruency factor. As in the first experiment, the congruency factor improved the model fit significantly for solving times [$\chi^2(1) = 4.27$, $p = .039$, $\beta = -218.01$, $t = -2.11$, CI (-420.33, -15.58)]. The mean solving time for anagrams in matching trials was again faster compared to mismatching trials ($M_{\text{Match}} = 3.305$ s, $M_{\text{Mismatch}} = 3.537$ s). Detailed solving times, sorted by color of the prime and color of the solution word's referent can be found in Table 2. Looking at the reaction times from anagram onset to confirmed response again did not change this result [6171 trials under 20 s, compared to 6278 trials in the original analysis, $\chi^2(1) = 3.98$, $p = .046$, $\beta = -185.97$, $t = -2.03$, CI ($-365.81.33$, -6.14)]. Just like in the first experiment, the congruency factor did not influence the accuracy rates [$\chi^2(1) = 2.53$, $p = .28$, $\beta = 0.004$, $t = 0.48$, CI (-0.014, 0.023)], with mean accuracy rates of 80.31% in matching trials and 79.78% in mismatching trials. Including list as an additional factor did not further improve the fit for the prediction of solving times or accuracy (both $ps > 0.4$).

Since we were interested whether there was a difference between modal and amodal cues, we pooled the data for both experiments and used another mixed models analysis with the same null model as was used for the individual experiments, but added experiment as a factor. While the congruency effect remained robust [$\chi^2(1) = 11.62$, $p < .001$, $\beta = -257.40$, $t = -3.57$, CI (-398.91, -115.89)] compared to the null model, neither a model with an added factor for type of experiment [$\chi^2(1) = 0.46$, $p = .50$, $\beta = -93.15$, $t = -0.68$, CI (-361.70, 175.40)] nor a model with an interaction between congruency and experiment [$\chi^2(2) = 0.61$, $p = 0.74$, congruency: $\beta = -230.51$, $t = -2.30$, CI (-426.80, -34.21), type of experiment: $\beta = -73.47$, $t = -0.50$, CI (-359.87, 212.92), congruency * experiment: $\beta = -53.48$, $t = -0.39$, CI (-323.93, 216.98)] provided a better data fit than the model with the factor congruency only.

Discussion

In our second experiment, we replicated the results from the first experiment despite using color patches instead of color words as cues. Given the amodal and abstract nature of the anagram solving task that was used here and the mixed previous results regarding cross-modal priming in anagram solving tasks, it is quite surprising that modal cues, like color patches, seem to have a similar influence on the anagram solving process as color words. This finding seems to support the ideas of embodied or situated accounts of cognition.

While there did not seem to be a difference between the two experiments, as the combined analysis has shown, we were interested in gaining more information on the influence of modal and amodal cues on conceptual representations. In Experiment 3, we combined the first two experiments, showing participants a color word inside a colored rectangle to find out whether the two cues canceled each other out when combined, or if either the amodal or modal cue would prove to be dominant over the other.

Experiment 3

In this experiment, we used amodal and modal cues prior to anagram solving. Participants were presented with color words inside of color patches that either both matched, both mismatched or partially matched the color that was associated with the anagram's solution word. If amodal cues are more important than modal cues for anagram solution processes, then we should observe faster solution times in conditions in which only the amodal cues match compared to the conditions in which only the modal cues match. The opposite response time pattern should be observed if modal cues are more important than amodal cues.

Participants

Participants were recruited by sending out a link via E-Mail to the remaining students of the University of Tübingen who had not received an invitation to one of the earlier experiments. Participation was voluntary, and participants could enter yet another lottery to win 1 of 6 20 € vouchers for an online shop. 145 participants completed the experiment (103 female, age ranging from 18 to 65 years, $M = 23.82$).

Material and procedure

The material as well as the procedure stayed mostly the same as in the first two experiments. Only this time, instead of a single color word or a colored patch, participants saw a color word written in black font color inside a colored rectangle. We used the same colored rectangles as in the second experiment and both cues were presented at the same time for 700 ms. To keep all factors balanced across participants without any participant seeing the same anagram twice, we expanded our lists from 4 to 8. Across these lists, every target word was shown once with each prime color and color word. This resulted in 20 congruent trials regarding prime color (the color of the rectangle matched the associated color of the solution word's referent) and 20 congruent trials regarding color word (the color word matched the associated color of the solution word's referent). In 20 of these 40 congruent trials, prime color and color word matched, leading to a double congruency (e.g. participants saw a green rectangle with the word "green" written in the middle of it followed by an anagram of "cucumber"). The remaining 50 trials in each list were incongruent trials, where neither color word nor prime color matched the associated color of the solution word's referent.

Design

The design for our third experiment was a 4 (prime color) × 4 (color word) × 4 (referent color) × 8 (list) design, with prime color, color word and referent color as within-participant variables and list as a between-participant variable.

Results

We used the same exclusion criteria as in the first two experiments: participants with less than 66% correct answers in the experimental trials were excluded (excluding eight participants). As before, the Ns of the lists were equalized, excluding the most recent participants from four of the eight lists to bring the number of participants down to the number of the lists with the least participants (excluding nine participants). This resulted in 128 participants (91 female, age ranging from 18 to 65, $M = 23.83$) who were included in the final analysis, 16 per list. Removing solving times above 20 s reduced our dataset by less than 6%. For the analysis of solving times only correct trials were taken into account.

We used a base model consisting of fixed main effects for anagram length and word frequency, random intercepts for participants and items as well as by-participant and by-item random slopes for prime color congruency * color word congruency. Neither adding the factor prime color congruency [$\chi^2(1) = 2.74$, $p = .098$, $\beta = -146.07$, $t = -1.68$, CI $(-316.75, 24.61)$] nor the factor color word congruency [$\chi^2(1) = 1.96$, $p = .161$, $\beta = -135.86$, $t = -1.42$, CI $(-323.49, 51.78)$] improved the fit of the model significantly. Also neither an additive model with both prime color congruency and color word congruency [$\chi^2(2) = 3.33$, $p = .190$, color word congruency: $\beta = -81.83$, $t = -0.78$, CI $(-288.90, 125.24)$, prime color congruency: $\beta = -113.32$, $t = -1.18$, CI $(-302.34, 75.70)$] nor a model with the interaction of both factors yielded a significant improvement [$\chi^2(3) = 4.31$, $p = .230$, color word congruency: $\beta = -2.35$, $t = -0.18$, CI $(-260.65, 255.95)$, prime color congruency: $\beta = -43.26$, $t = -0.36$, CI $(-277.75, 191.24)$, prime color * color word congruency: $\beta = -189.75$, $t = -1.00$, CI = $(-562.00, 182.50)$]. Adding list as a factor did not improve the fit of any of these models (all ps > 0.6).

In a post-hoc analysis, we looked at each kind of congruency separately, splitting our data set into prime color congruent (but not doubly congruent), with incongruent trials being color word congruent and doubly incongruent trials, color word congruent (but not doubly congruent) with incongruent trials being prime color congruent and doubly incongruent trials and doubly congruent with incongruent trials being prime color or color word congruent as well as doubly incongruent trials. Then, we used a base model consisting of fixed main effects for anagram length and word frequency, random intercepts for participants and items (but without by-item and by-participant random slopes, since not all models converged using these), and compared it to a model with the congruency factor added to it for each subset. While congruency did not improve the model fit significantly in the prime color congruent and color word congruent subsets [$\chi^2(1) = 0.13$, $p = .720$, $\beta = -40.24$, $t = -0.36$, CI $(-260.51, 180.02)$ and $\chi^2(1) = 0.07$, $p = .933$, $\beta = -9.446$, $t = 0.084$, CI $(-230.52, 211.62)$, respectively], it yielded a significantly better fit for the doubly congruent subset [$\chi^2(1) = 4.41$, $p = .036$, $\beta = -231.79$, $t = -2.10$, CI $(-448.16, -15.43)$]. Adding the time for entering the response to the solving times once again did not change these results [prime color congruent: 6003 trials under 20 s compared to 6071 trials in the original analysis, $\chi^2(1) = 0.15$, $p = .701$, $\beta = 39.60$, $t = 0.38$, CI $(-162.88, 242.08)$, color word congruent: 6005 trials under 20 s compared to 6086 in the original analysis, $\chi^2(1) = 0.44$, $p = .507$, $\beta = -68.25$, $t = -0.66$, CI $(-269.73, 133.23)$, doubly congruent: 6017 trials under 20 s compared to 6089 trials in the original analysis, $\chi^2(1) = 5.14$, $p = .023$, $\beta = -227.94$, $t = -2.27$, CI $(-424.94, -30.95)$].

Table 3 Mean solving times in ms and percentage of correctly solved anagrams for each type of congruency

Congruency	Mean solving time	Percentage correct
Incongruent	3391	82.65
Prime color	3350	80.88
Color word	3346	82.88
Double congruency	3140	82.78

For the accuracy rates, we had to cut the by-item and by-participant random slopes for color word * prime color congruency since the models did not converge otherwise. Similar to the reaction time analysis, neither prime color congruency [$\chi^2(1) = 2.14$, $p = .144$, $\beta = -1.10$, $t = -1.46$, CI $(-2.57, 0.37)$] nor color word congruency [$\chi^2(1) = 0.96$, $p = .33$, $\beta = 0.73$, $t = 0.98$, CI $(-0.73, 2.20)$] were improvements over the null model. An additive model with both prime color congruency as well as color word congruency [$\chi^2(2) = 4.70$, $p = .10$, color word congruency: $\beta = 1.28$, $t = 1.60$, CI $(-0.29, 2.85)$, prime color congruency: $\beta = -1.55$, $t = -1.93$, CI $(-3.12, 0.02)$] and a model with the interaction of both factors [$\chi^2(3) = 5.54$, $p = .136$, color word congruency: $\beta = 0.72$, $t = 0.72$, CI $(-1.25, 2.70)$, prime color congruency: $\beta = -2.11$, $t = -2.10$, CI $(-4.09, 0.14)$, prime color * color word congruency: $\beta = 1.50$, $t = 0.92$, CI $(-1.69, 4.69)$] still yielded no significant improvement. Adding list as a factor did not improve the fit of any of these models (all ps > 0.2). A summary of the mean solving times and the percentage of correctly solved anagrams in the different congruency conditions is found in Table 3.

Discussion

The results showed that neither a matching amodal nor modal color cue facilitated anagram solving when the other cue was mismatching. Only when both types of cues matched the associated color of the solution word's referent, anagram solving was facilitated. It can be concluded that the cuing effects we observed in the first two experiments both seem to be rather fragile. When combined, neither modal nor amodal cues dominate over the other. Instead, they both seem to activate the same superimposed conceptual color representation that is only helpful in solving anagrams when no other conceptual color representation is active at the same time.

General discussion

We conducted a series of anagram solving tasks using anagrams of nouns that refer to entities that are associated with one of four colors. Before each anagram either a color word (Experiment 1) or a colored rectangle (Experiment 2) was presented, that either matched or mismatched the associated color of the solution word. We found faster solving times in trials where the color cue matched the associated color of the solution word compared to mismatching trials in both experiments. In Experiment 3, participants saw both the amodal (color word) as well as the modal (colored rectangle) cue at the same time. A facilitation was only present when both color cues matched the color of the solution word's referent.

Our results may serve as another piece of evidence that multimodal conceptual representations play an important role in language comprehension. Such conceptual representations evoked by language have been shown to involve experiential traces pertaining to different meaning dimensions, such as shape (Zwaan, Stanfield, & Yaxley, 2002), orientation (Stanfield & Zwaan, 2001), vertical position (Lachmair et al., 2011), and color (Connell, 2007; Zwaan & Pecher, 2012; Mannaert et al., 2017). Berndt et al. (2016) were able to show the corresponding influence in the reverse direction by showing anagrams either at the top or at the bottom of the screen, serving as a prime of vertical position. They found faster solving times when the anagram's position on the screen matched the associated vertical position of the solution word, but only when additional contextual information was provided (in the form of a background picture matching the theme of the solution words).

In the current study, we expanded upon this finding, using color as a prime to activate concepts associated with it, thus further showing that modal experiential traces can activate conceptual representations and the words associated with them in the same way as amodal cues do. Interestingly, color as a contextual modal cue was enough to activate words associated with that color. No additional contextual cues were needed, as was the case in the original anagram solving study by Berndt et al. (2016). This may be due to the fact that color is an intrinsic object feature that also plays an important role in object recognition (Bramão, Reis, Petersson, & Faísca, 2011) whereas an object's typical position in vertical space does not refer to an intrinsic object feature but rather relates the object in question to other objects in the surrounding world and may thus be less easily accessible. To test this explanation of the differences between the results of the current and our previous study, additional priming studies with an anagram solution task looking at other intrinsic and extrinsic object properties would be needed.

Richter and Zwaan (2010) have already shown that color and shape cues can activate conceptual representations facilitating access to associated words in a categorization, a lexical decision, and a naming task. They also investigated how both cues were integrated by showing both a matching or mismatching shape and a matching or mismatching color. If both cues were integrated additively both cues should have contributed their individual match effect, leading to moderate match effects if either cue matched and a stronger match effect when both cues matched. What they found instead was that both cues were integrated multiplicatively: when both shape and color matched the target, the resulting match effect was much larger than what the individual match effects of color and shape would have suggested. Their results are similar to what we observed in our third experiment. It is important to note, however, that we used a rectangular shape as our color cue and did not control for the shape of the solution words referents, thus it is unclear whether the match effect for color that we observed was actually weakened in most trials because of a mismatching shape cue.

As outlined in the introduction there is one major dichotomy regarding the way meaning is represented. On the one side, propositional models argue for a fully symbolic, amodal representation of meaning (e.g., Fodor, 1975) on the other side embodied models of language comprehension suggest that meaning is represented in a modal way that directly resembles sensorimotor experiences. Interestingly, recently various hybrid models of language representation have been proposed. For example, Binder and Desai (2011) proposed that concepts are stored in convergence zones between the different sensory and motor areas in the brain. This assumption also implies cascading activation to sensory and motor areas depending on the context and the task, but the actual conceptual representation are supposed to be supramodal, storing information about cross-modal conjunctions: "conceptual representation consists of multiple levels of abstraction from sensory, motor and affective input. […] The top level contains schematic representations that are highly abstracted from detailed representations in the primary perceptual-motor systems". (Binder & Desai, 2011, p. 531). This view is a hybrid approach between traditional and embodied accounts of language comprehension, including both abstract and multimodal representations (also see Dove, 2009, 2011). One now could speculate that our

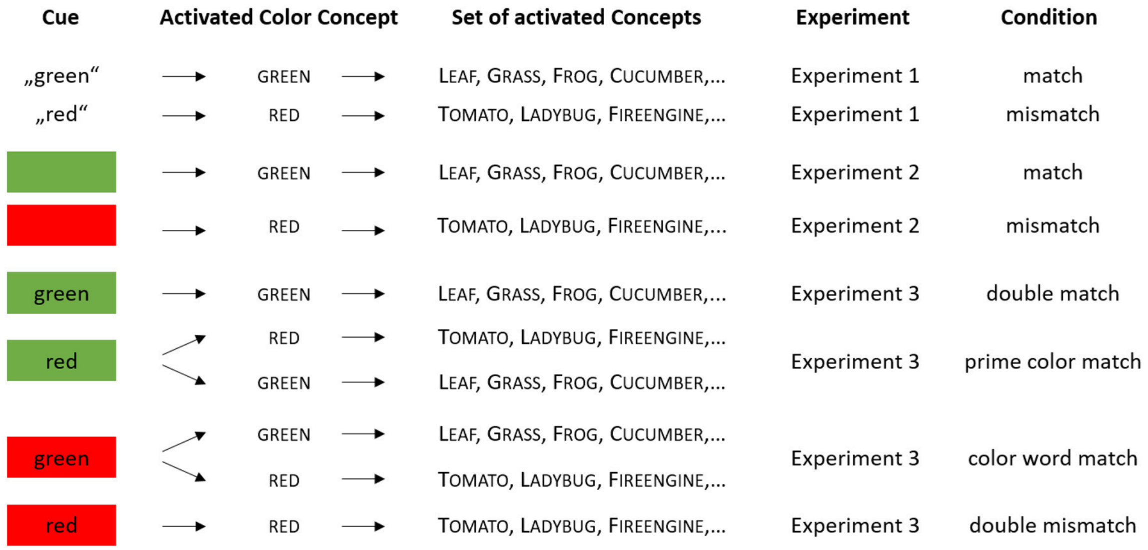


Fig. 1 Amodal detour account explaining the results of Experiments 1–3 (illustrating the different conditions for "cucumber" as the solution word). The account assumes that a modal color cue activates the corresponding amodal color concept which in turn then activates a set of concepts that are associated with this color. Anagram solution is facilitated if the solution word is part of a rather small set of activated concepts (as in the case of the match conditions of Experiments 1 and 2 and the double match condition of Experiment 3). If the solution word is not activated (as in the mismatch conditions of Experiments 1 and 2) or only activated as part of a larger set of concepts (as is the case in the prime color match, the color word match, and the double mismatch conditions in Experiment 3), no facilitation occurs. (Color figure online)

results are in line with hybrid accounts of cognition, where conceptual representations are strongly interconnected with sensory, motor and affective systems but are not limited to these multimodal traces. Proponents of such a hybrid view of meaning representation might argue that if conceptual representations were exclusively multimodal, like strong embodiment accounts suggest (e.g., Gallese & Lakoff, 2005), one would have predicted a stronger match effect of the modal color cue compared to the amodal color word. The same logic may also apply to strong disembodiment theories (e.g., Fodor, 1975), where one would have expected a stronger match effect of the amodal color cue. However, proponents of non-hybrid views of meaning representation may assume that modal/amodal color cues are effective only by means of themselves activating another cue (amodal or modal, respectively) that then directly influences anagram solution processes. Indeed, Mahon and Caramazza argued for such a disembodied view of cognition, where the "activation cascades from disembodied concepts to sensory and motor systems that interface with the conceptual system" (p. 60, 2008). According to them, the activation of sensory and motor areas is merely the result of a dynamic flow of activation between and within cognitive systems, but originating from abstract conceptual representations. The idea that multimodal activation is merely a result of cascading activation may also seem implausible given our results. When both a modal and an amodal cue activate conflicting amodal color concepts, but the modal cue does so only by a "detour" of cascading back to the amodal representation, one would have expected a stronger match effect for amodal cues since these should have activated the corresponding color concept earlier and more directly. Thus, proponents of a hybrid view of meaning representation may take the fact that neither the amodal nor the modal color cue proved to be dominant in our studies as evidence against the idea of non-hybrid "detour accounts".

However, it should be taken into account that anagram solving times are quite long. The reasoning in the previous paragraph would be more plausible if our study had involved a task with rather short reaction times. In a complex task such as anagram solving, the time required for a potential detour (e.g., a modal cue activating the corresponding color word) may simply not be long enough to have a significant impact on the overall solving times. In other words, in our view, the present results do not allow ruling out non-hybrid detour accounts and thus do not unequivocally speak in favor of hybrid forms of meaning representation. Figure 1 illustrates one version of an amodal detour account, according to which modal cues are effective only by means of activating amodal color concepts (e.g., by means of verbalization processes). These amodal color concepts then in turn prime concepts that are associated with that color. Anagram solution processes are facilitated if the activated color concept primes a rather small set of concepts including the solution word. Conditions in which conflicting cues active several color concepts do not sufficiently activate the solution word because the set of activated concepts is too large. It is noteworthy, however, that Berndt et al. (2016) conducted a control experiment, telling participants that solution words will be associated with either the ocean or the sky and it turned out that a smaller set of activated concepts alone did not sufficiently explain their results in the previous experiments. In principle it also seems possible that conditions with conflicting cues may fail to sufficiently activate the solution word because the presence of a color word in the cue that is not identical to the color of the background patch (e.g., the word "red" on a green color patch) may interfere with verbalizing processes such that the solution concept is not pre-activated at all. However, it should be noted that this version of the detour account cannot explain why the color word match condition of Experiment 3 did not lead to facilitated anagram solution processes. After all, in this condition, the matching color word should have pre-activated the solution word without interference from the inconsistent color patch.

To investigate the viability of detour accounts in more detail, additional studies are required, some of which we are currently working on.[1] In one experiment, the anagrams themselves are presented in different font colors, instead of providing a color cue before showing the anagram. By changing the color of the anagram from black to another font color, we can modulate the onset of the modal color cue exploring its influence in different stages of anagram solving. In another experiment, we directly test the role of verbalization processes. To make sure that participants cannot verbalize the color of the color patch, we have participants say a string of nonsense words while seeing a modal color cue. If the congruency effect we observed in our second experiment was only due to participants verbalizing the color of the patch (see account in Fig. 1), we should not find any congruency effects with this experimental setup. If we still find an influence of the color patch however, it would seem unlikely that verbalization processes underlie the congruency effects observed in our experiments.

In addition to providing important insights into the nature of the meaning representations utilized in language comprehension, using anagram solving tasks with different contextual cues could also prove helpful to study the mechanisms involved in problem solving. Associative priming has been shown to facilitate anagram solving by presenting other words like "chair" as a prime for "table" (Dominowski & Ekstrand, 1967) and seems to work even if it is provided in the form of modal information associated with the solution words (in this case color). Thus, it seems that problem solving does not solely operate on abstract symbolic representations as proposed by classical problem solving theories (e.g., Newell & Simon, 1972), but can also process information from at least the visual modality effectively. Further studies could investigate whether olfactory or auditory information can be used in a similar way to support the problem solving process, and if so, whether such a beneficial effect of associative priming is constrained to language or extends to problem solving in general. Kirsh mentioned the importance of a theory of hints for problem solving and stated that "a theory of situated problem solving should explain why hints are successful and the many ways our environments offer us hints on how to solve our problems" (p. 293, 2009). Knowing the basis of our conceptual representations, how they can help in problem solving and how they are activated voluntarily or automatically is an important step in this direction.

One more aspect to keep in mind is that there may be different approaches to solving anagrams, and that some participants may even be able to choose between different strategies depending on the task at hand. Novick and Shermann (2003, 2008) have indeed observed different strategies between skilled and less skilled anagram solvers. We cannot exclude the possibility that contextual cues may only benefit one of those strategies but not the other. In this vein, it is important to keep in mind that our participants were mainly students of the University of Tübingen and thus represent only a small, well-educated subsample of the general population (for a critical discussion, see Jones, 2010). The current study was successfully conducted as a browser-based experiment that did not require participants to come to a laboratory. Future browser-based studies could thus be used to reach a bigger and more divergent pool of participants, not restricted to university students, focusing more on interindividual differences and their influence on anagram solution strategies.

Conclusion

Presenting a modal cue prior to an anagram (a color patch) can facilitate solution processes when it matches properties of the solution word's referent (e.g., a red patch for an anagram of tomato) in the same way as an amodal cue (a color word) does. When combined, neither cue shows a stronger match effect than the other, instead they cancel each other out. Only when both color cues activate the same color concept and when it is associated with the referent of the anagrams solution word, a facilitation is observed. Further research is ongoing to address unresolved theoretical issues.

Acknowledgements This study was supported by the German Research Foundation (DFG) as part of the Collaborative Research Center "SFB 833"; Project B4 of Barbara Kaup and Hartmut Leuthold. We want to thank the two reviewers of an earlier version of this paper for their helpful remarks and great suggestions.

---

[1] We thank Stephen Reed for suggesting experiments along these lines.

References

Antonietti, A., & Girotti, G. (1991). Anagram solving is facilitated by intra-but not cross-modal priming. *Perceptual and Motor Skills, 72*(2), 403–408.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences, 22*(4), 637–660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2015). lme4: Linear mixed-effects models using Eigen and S4, 2014. *R package version, 1*(4), 1–23.

Berndt, E., Dudschig, C., & Kaup, B. (2016). Activating concepts by activating experiential traces: Investigations with a series of anagram solution tasks. *The Quarterly Journal of Experimental Psychology, 71*(2), 483–498.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*(11), 527–536.

Bramão, I., Reis, A., Petersson, K. M., & Faísca, L. (2011). The role of color information on object recognition: A review and meta-analysis. *Acta Psychologica, 138*(1), 244–253.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition, 102*(3), 476–485.

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods, 48*(1), 1–12.

Dominowski, R. L., & Ekstrand, B. R. (1967). Direct and associative priming in anagram solving. *Journal of Experimental Psychology*, *74*(1), 84–86.

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition, 110*(3), 412–431.

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology, 1*, 242.

Dudschig, C., de la Vega, I., De Filippis, M., & Kaup, B. (2014). Language and vertical space: On the automaticity of language action interconnections. *Cortex, 58*, 151–160.

Dudschig, C., & Kaup, B. (2017). Is it all task-specific? The role of binary responses, verbal mediation, and saliency for eliciting language-space associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(2), 259–270.

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading “sun” and looking up: The influence of language on saccadic eye movements in the vertical dimension. *PloS One, 8*(2), e56872.

Dudschig, C., de la Vega, I., & Kaup, B. (2015). To fly or not to fly? The automatic influence of negation on language–space associations. *Cognitive Processing*, *16*(1), 203–207.

Fodor, J. A. (1975). *The language of thought*, vol. 5. Cambridge: Harvard University Press.

Gallese, V., & Lakoff, G. (2005). The brain’s concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3–4), 455–479.

Jones, D. (2010). A WEIRD view of human nature skews psychologists’ studies. *Science, 328*(5986), 1627–1627.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction- integration model. *Psychological Review, 95*(2), 163–182.

Kaup, B., de la Vega, I., Strozyk, J., & Dudschig, C. (2015). The role of sensorimotor processes in meaning composition. In M. H. Fischer & Y. Coello (Eds.), *Conceptual and Interactive Embodiment* (pp. 58–82). London: Routledge.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*(2), 173–202.

Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins, M. Aydede (eds.) *The Cambridge Handbook of Situated Cognition* (pp. 264–306). Cambridge, MA: Cambridge University Press.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychonomic Bulletin & Review, 18*(6), 1180–1188.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102*(1–3), 59–70.

Mannaert, L. N. H., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory* & *Cognition, 45*, 974–982.

Mehta, R., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science, 323*(5918), 1226–1229.

Mills, M., Boychuk, P., Chasteen, A. L., & Pratt, J. (2018). Attention goes both ways: Shifting attention influences lexical decisions. *Journal of Experimental Psychology: General, 147*(2), 282–291.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.

Novick, L. R., & Sherman, S. J. (2003). On the nature of insight solutions: Evidence from skill differences in anagram solution. *The Quarterly Journal of Experimental Psychology: Section A, 56*(2), 351–382.

Novick, L. R., & Sherman, S. J. (2008). The effects of superficial and structural information on online problem solving for good versus poor anagram solver. *The Quarterly Journal of Experimental Psychology, 61*(7), 1098–1120.

Rajaram, S., & Roediger, H. L. (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(4), 765–776.

Richter, T., & Zwaan, R. A. (2010). Integration of perceptual information in word access. *Quarterly Journal of Experimental Psychology, 63*(1), 81–107.

Scribner, S. (1985). Knowledge at work. *Anthropology & Education Quarterly, 16*(3), 199–206.

Sperber, R. D., McCauley, C., Ragain, R. D., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory & Cognition, 7*(5), 339–345.

Srinivas, K., & Roediger, I. I. I., H. L (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language, 29*(4), 389–412.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153–156.

Steele, K. M. (2014). Failure to replicate the Mehta and Zhu (2009) color-priming effect on anagram solution times. *Psychonomic Bulletin & Review, 21*(3), 771–776.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS One, 7*(12), e51382.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.

C. Study 3:

Copyright Notice

Berndt, E., Dudschig, C., Miller, J., & Kaup, B. (2019). A replication attempt of hemispheric differences in semantic-relatedness judgments (Zwaan & Yaxley, 2003). *Acta Psychologica*, *198*, 102871. https://doi.org/10.1016/j.actpsy.2019.102871

Acta Psychologica 198 (2019) 102871

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

A replication attempt of hemispheric differences in semantic-relatedness judgments (Zwaan & Yaxley, 2003)

Eduard Berndt[a,*], Carolin Dudschig[a], Jeff Miller[b], Barbara Kaup[a]

[a] Department of Psychology, University of Tübingen, Tübingen, Germany
[b] Department of Psychology, University of Otago, Dunedin, New Zealand

ARTICLE INFO

Keywords:
Language representation
Hemispheric differences
Embodied cognition

ABSTRACT

In a study by Zwaan and Yaxley (2003, *Cognition*, *87*, B79–B86), participants judged the semantic relatedness of word pairs presented one above the other either in the left or right visual field with all related pairs requiring right-handed responses. If the vertical orientation of the word pairs matched their referents' typical vertical orientation ("*roof*" above "*basement*") a match effect was observed, but only when the word pair was presented in the left visual field. We replicated this study with response side as an additional factor and found a main effect of match, as well as a Simon effect with faster responses when the required response matched the visual field in which the word pair was presented. We did not, however, observe an interaction between the match effect and the visual field. This challenges the assumption that coarse semantic representations, including spatial properties of objects, are mainly processed in the right hemisphere.

1. Introduction

Perceptual theories of mental representation have gained a lot of support in the past two decades. In contrast to more traditional abstract theories of mental representation (e.g., Fodor, 1975; Pylyshyn, 1986), these theories assume that mental representations are sensorimotor in nature even when higher cognitive processes are involved (Barsalou, 1999, 2008). For instance, for language comprehension, proponents of perceptual theories assume that upon reading or hearing a sentence or a word, past sensorimotor experiences (e.g., visual, olfactory) related to the meaning of the word are re-activated and eventually give rise to a simulation of the meaning of what was read or heard (Zwaan & Madden, 2005). Indeed, there have been several studies showing the involvement of sensorimotor processes during language comprehension. For instance, spatial information stemming from experiencing the referents of linguistic expressions has been shown to be reactivated during comprehension. Corresponding results have been reported in both perceptual tasks (e.g., Estes, Verges, & Barsalou, 2008; Meteyard, Bahrami, & Vigliocco, 2007; Ostarek & Vigliocco, 2017), as well as motor tasks (Dudschig & Kaup, 2017; Dudschig, Lachmair, de la Vega, De Filippis, & Kaup, 2012; Lachmair, Dudschig, De Filippis, de la Vega, & Kaup, 2011). For example, responses to words such as "bird" with a typical referent location in the upper visual field are faster when they involve an upwards movement with the arm compared to a downwards movement. The opposite holds for words such as "shoe" with a typical referent location in the lower visual field. It is proposed that such associations between words and the spatial aspects of non-linguistic experiences already originate during language learning. More specifically, during learning words and referents are typically co-present such that the two experiences get associated. Later when only the word is being processed, the corresponding sensorimotor experiences are automatically activated. On a neurobiological level this learning mechanism is often explained via basic Hebbian learning principles (Pulvermüller & Fadiga, 2010; Pulvermüller, 2013, 2018; Zwaan & Madden, 2005; for an overview see also Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Fischer & Zwaan, 2008).

Thus, the simulations created during comprehension are assumed to contain information regarding sensory aspects of the word's referents. Some authors also emphasize the idea that simulations created during comprehension contain information regarding the relationships between the mentioned referents. Previous research has indeed investigated the role of spatial relationships between referents, with the assumption being that spatial relations are captured in an analogue manner in the mental simulations created during language comprehension (e.g., Glenberg, Meyer, & Lindem, 1987; Morrow, Bower, & Greenspan, 1989). Of particular importance to the current study are two studies by Zwaan and Yaxley (2003a, 2003b) in which the authors observed a match effect between the spatial arrangement of two words

---

[*] Corresponding author at: Schleichstrasse 4, 72076 Tübingen, Germany.
E-mail address: eduard.berndt@uni-tuebingen.de (E. Berndt).

https://doi.org/10.1016/j.actpsy.2019.102871
Received 21 December 2018; Received in revised form 4 June 2019; Accepted 10 June 2019

presented on a computer screen and the canonical spatial relationship between the referents of these words: A word like *"nose"* written above *"mustache"* (matching the referents' spatial relation to each other) led to faster responses in a semantic-relatedness-judgment task compared to *"nose"* appearing below *"mustache"* (mismatching the referents' spatial relation to each other). The authors concluded that participants engaged in perceptual simulations of their referents when presented with the word pairs, and that these perceptual simulations influenced the response times in the semantic-relatedness judgments. In particular, Zwaan and Yaxley (2003b) state: *Take for example noses and moustaches. They are specific parts of faces. Langacker, 1998 has suggested that in cases such as these, perceptual representations of the larger object of which the constituents are parts will be activated, in this example a face, with the focus on the part denoted by the noun in question* (e.g. *nose). This process of mental simulation will make the iconic match or mismatch between the visual representation of the words on the screen and their referents in the mental simulation available to the subject. The results of this perceptual simulation in turn influenced the semantic-relatedness judgments, such that these were faster in the case of a match than in the case of a mismatch.* (p. 80).

Zwaan and Yaxley (2003b) used a divided visual field paradigm (DVF; see Bourne, 2006) to investigate potential differences between the two hemispheres regarding perceptual simulations during language processing. By presenting stimuli very briefly to either the left visual field (LVF) or the right visual field (RVF), the stimulus is initially only received and processed by the hemisphere contralateral to its presentation side. Finding an interaction between hemisphere and match, with the match effect being confined to the RH, the authors concluded that perceptual simulation of the words' referents takes place predominantly in the RH, which is in line with the assumption that closely related semantic associations are activated in the LH upon reading or hearing a word, while more distant, peripherally related semantic associations are processed in the RH (Beeman, 1998; Beeman & Bowden, 2000).

However, there is an alternative explanation for these results, because the authors did not counterbalance the response side. All experimental word pairs required a yes-response and this was always given by pressing a button on the ride side with the right hand. Thus, the condition in which a match effect was observed was a condition in which the stimuli appeared at the side contralateral to the required response. It is well known that response latencies are strongly affected by whether the stimuli are presented at the same or the opposite side as the required response. This is the Simon effect (Simon & Rudell, 1967), describing the phenomenon that responses tend to be faster and more accurate when the stimulus and response sides match than when they mismatch. In accordance with this phenomenon, Zwaan and Yaxley (2003b) indeed found slower responses to word pairs presented to the LVF compared to word pairs presented to the RVF. In principle, it is possible that the observed differences in the two visual field conditions do not reflect hemispheric differences but rather differences that are due to the relationship between stimulus location and response side, with a match effect being observed only when stimuli are presented contralateral to the required response. To investigate the possibility that the match by hemisphere interaction observed by Zwaan and Yaxley (2003b) does not reflect hemispheric differences in perceptual simulations but rather differences that are due to an interaction between stimulus location and response side, we decided to conduct a replication study of the Zwaan and Yaxley study and to manipulate response side as a between-participants factor.

For the current study, we expected to find a Simon effect, resulting in an interaction between visual field (VF) and response side. Since we only added response side as a factor but kept the rest of the design the same, we expected to replicate the interaction between VF and match, as reported by Zwaan and Yaxley (2003b). If there are indeed hemispheric differences in processing word pairs with a canonical spatial relation, we would expect these differences to persist with left-handed responses. In other words, we should observe an interaction between match and VF, averaging across response sides. In contrast, if the observed differences did not reflect hemispheric differences but were rather due to an interaction between the match effect and the Simon effect (match effect only for the side contralateral to the required response), then we might observe a three-way interaction between match, VF, and response side. In other words, we would replicate the results by Zwaan and Yaxley for right-hand responses (i.e., a match effect for the left VF but not for the right VF) but observe the opposite pattern for left-hand responses (i.e., a match effect for the right VF but not for the left VF). In any case, if we do not find an interaction between match and VF, we can conclude that the spatial match effect is not confined to either one of the hemispheres.

2. Method

2.1. Participants

One-hundred-eighty-three participants were recruited among interested students of the University of Tübingen for money or course credit. Applying the same exclusion criteria as Zwaan and Yaxley (2003b), we excluded participants with less than 50% correct responses in one of the conditions (54 participants equalling 29.51% of all participants) as well as participants with less than 70% correct responses overall (another 22 participants equalling another 12.02% of all participants). Similar to the original study, this resulted in the exclusion of about half of the participants which was expected considering the very short presentation time of the stimuli. Another 11 participants were excluded to equalize the number of participants in each list. Ninety-six right-handed participants were included in the final analysis of the experiment (67 female, age ranging from 18 to 36, $M = 22.4$). All participants provided informed written consent.

2.2. Material

The stimulus material consisted of the 128 word pairs with concrete nouns used in the original study by Zwaan and Yaxley (2003b), translated to German (e.g. ROOF-HOUSE was translated to DACH-HAUS). Forty-four of these word pairs described objects that have a fixed associated vertical position in relation to the other object of that word pair. For example, the ceiling is always located above the floor in a canonical view of a room. All of these items were semantically related. The number of letters of the words in these critical word pairs ranged from four to twelve ($M = 6.34$; note that it ranged from four to ten, with $M = 5.73$ in the original study using English words). In addition to the experimental items, 84 filler word pairs were used, of which 20 were semantically related while the other 64 were not. These filler items did not mention referents in typical vertical orientation relative to each other. Since we used the same items as Zwaan and Yaxley (2003b) and they had already tested their items for semantic relatedness using LSA (Landauer & Dumais, 1997), we did not test for semantic relatedness again.

2.3. Procedure

The experiment was conducted with *E*-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) using a template of the original study from Zwaan and Yaxley (2003b)[1] on a computer with a 120 Hz cathode ray tube monitor. The stimuli appeared in four invisible boxes on the screen (upper left, lower left, upper right and lower right) with both words of a word pair appearing either in the LVF or the RVF. A chin rest with a distance of 51 cm to the screen was used to make sure

---

[1] Many thanks to Richard Yaxley for providing us with the original *E*-Prime files of the experiment.

that the viewing distance was the same across all participants and trials. The horizontal center of the words was 2.7 cm (3° visual angle) left or right from the center of the screen, while the vertical center of the words was 0.6 cm (0.69° visual angle) above or below the center of the screen. The inner edge of the shortest words (four letters) had a distance of 1.9 cm (2.13° visual angle) to the center of the screen while it was 3.7 cm (4.16° visual angle) for the outer edge. The inner edge of the longest words (twelve letters) was at the center of the screen, while the outer edge had a distance of 5.7 cm (6.40° visual angle) to the center of the screen.

Participants were instructed to read the word pairs that would appear on the screen very briefly and to judge whether they were semantically related or not by pressing either the F-key or the J-key on the keyboard with the left or right index finger, respectively. While yes-responses were always assigned to the J-key in the original study by Zwaan and Yaxley (2003b), we counterbalanced the assignment across participants in our replication attempt. One half of the participants pressed the J-key for yes-responses and the F-key for no-responses, and the other half pressed the F-key for yes-responses and the J-key for no-responses. After presentation of the instructions, a short practice block consisting of four trials started, so that participants could familiarize themselves with the task before the experiment proper. At the beginning of each trial, a fixation cross was presented for 500 ms, followed by the presentation of the word pair for 200 ms. This brief presentation time was necessary to ensure that participants were not able to re-fixate the words after a saccade (see Zwaan & Yaxley, 2003b). To stick to the original procedure by Zwaan and Yaxley, we did not present a mask after word pair offset. The trial was logged as incorrect if no response was given within two seconds after stimulus presentation.

2.4. Design

We implemented four different lists, with each word pair being shown in both visual fields (LVF and RVF) across the lists. Every participant saw only one of the lists and thus saw each word pair only once in either LVF or RVF. Additionally, we varied which word of the word pair appeared above the other across lists. This resulted in a match-factor for our critical word pairs, with "match" meaning the words appeared on the screen according to the canonical spatial relationship of their referents (e.g., *"nose"* appearing above *"mustache"*) and "mismatch" meaning the words appeared on the screen in a spatial relationship that was opposite to the canonical spatial relationship of their referents (e.g., *"mustache"* appearing above *"nose"*). The response side that was required for a yes-response was manipulated between participants.

Our experiment thus implemented a 2 (VF: left vs right) × 2 (match vs mismatch) × 2 (response side: J-key as yes-response vs F-key as yes-response) × 4 (list) design, with VF and match as within-participant variables and response side and list as between-participant variables. Reaction times and error rates were our dependent variables.

Table 1
Mean latencies and accuracy rates of Experiment 1 segregated by match, visual field and response side with standard deviations in parentheses.

	Right handed responses			
	RVF (LH)		LVF (RH)	
	Latency	Accuracy	Latency	Accuracy
Match	828 (265)	0.83 (0.13)	864 (250)	0.81 (0.11)
Mismatch	836 (254)	0.84 (0.11)	874 (255)	0.82 (0.15)

	Left handed responses			
	RVF (LH)		LVF (RH)	
	Latency	Accuracy	Latency	Accuracy
Match	812 (252)	0.79 (0.13)	810 (247)	0.83 (0.13)
Mismatch	820 (253)	0.81 (0.13)	831 (255)	0.84 (0.13)

3. Results

Before analyzing the data, we applied the same exclusion criteria that were used in the original study by Zwaan and Yaxley (2003b), excluding participants with less than 50% correct trials in one or more of the 2 VF × 2 match/mismatch conditions or with less than 70% correct trials overall as has been reported in the participants subsection. One of the word pairs had to be excluded due to a programming error. To replicate the exclusion criteria of Zwaan and Yaxley (2003b), we also excluded reaction times below 200 and above 1700 ms (excluding 6.66% of the data), as well as trials with an RT that exceeded a distance of two standard deviations from a participant's mean (excluding 4.45% of the data).

We analyzed our data by means of linear mixed effect models (Baayen, Davidson, & Bates, 2008), but because Zwaan and Yaxley (2003b) analyzed their data by means of a by-subject and a by-item ANOVA in their original study, we will also present the results of these analyses in a footnote below.[2] Mean reaction times as well as accuracy rates across all conditions are summarized in Table 1 and visualized in Fig. 1. The left side of this figure shows the mean reaction times in the four VF × match conditions and illustrates that numerically reaction times were shorter in the match compared to the mismatch condition in both VFs. The right side of this figure shows the mean reaction times in the four response side × VF conditions and illustrates that numerically mean reaction times were shorter when the response side matched rather than mismatched the VF (right response and RVF; left response and LVF compared to right response and LVF; left response and RVF).

We analyzed these data by using the R-package lme4 (Bates, Mächler, Bolker, & Walker, 2015). To find an appropriate random effects structure for the models - balancing Type I error probability and power - we followed the approach suggested by Matuschek, Kliegl, Vasishth, Baayen, and Bates (2017). Specifically, we first performed Likelihood-Ratio Tests on models with four different random effects structures containing random intercepts for subjects and items but varying in the complexity of the random slopes, namely a maximal model with by-item as well as by-subject random slopes, a model with only by-item but no by-subject random slopes, a model with only by-subject but no by-item random slopes and finally a model without by-item or by-subject random slopes (see Matuschek et al., 2017, for more details on this procedure). The model without any random slopes proved to be the most efficient model and thus our base model going forward consisted of a fixed effect for the number of letters of the word pair and random intercepts for participants and items. In a stepwise procedure (see Table 2), we added terms to find the model with the best fit to our data: We first compared this base model to models with one of the main effects added as fixed effects (response side, VF, and match respectively). Both the model with visual field as well as the model with match added as fixed effects proved to be superior to the base model in a likelihood ratio test ($\chi^2(1) = 6.92$, $p = .009$, $AIC = 49{,}974$ and $\chi^2(1) = 4.27$, $p = .039$, $AIC = 49{,}977$ respectively). Going forward,

---

[2] In addition to the mixed models we also analyzed the data with the ANOVAs that were used by Zwaan and Yaxley (2003b) with the added factor Response Side. For the F1 analysis, we found a main effect of Visual Field (F (1, 94) = 7.37, $p = .008$, $\eta_p^2 = 0.073$), and an interaction between Visual Field and Response Side (F (1,94) = 4.65, $p = .034$, $\eta_p^2 = 0.047$). The main effect of Match was not significant (F (1,94) = 3.29, $p = .073$, $\eta_p^2 = 0.038$). The same held for the F2 analysis (main effect of Response side: F (1,42) = 25.22, $p < .001$, $\eta_p^2 = 0.375$; interaction between Response side and Visual Field: F (1,42) = 6.03, $p = .018$, $\eta_p^2 = 0.126$; main effect of Match: F (1,42) = 2.12, $p = .15$, $\eta_p^2 = 0.048$).

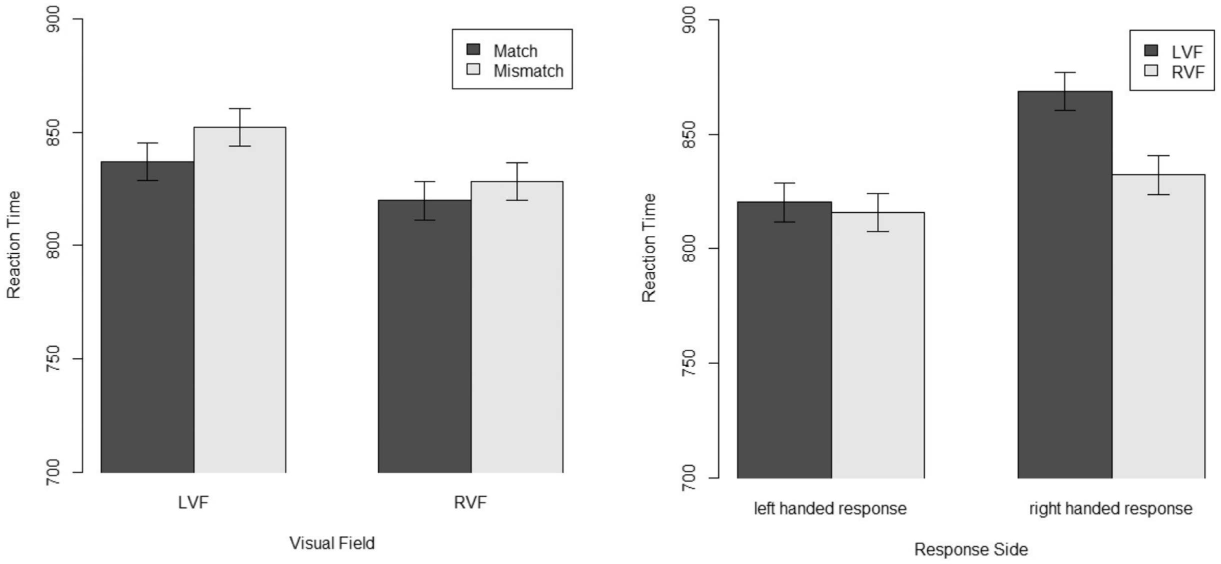


Fig. 1. Mean reaction times as a function of visual field and match condition are depicted on the left and mean reaction times as a function of response side and visual field are depicted on the right. Error bars represent standard errors.

Table 2

Overview of our stepwise procedure to determine the model with the best fit: The best model at each step is indicated by the bold font. The dashed lines separate hierarchical steps, meaning the best model from the previous step is tested against models with all possible combinations of the factors added, at first adding the main effects, later on also adding the interaction effects. Abbreviations: RT = Response Time, L = combined number of Letters of both words, VF = Visual Field, RS = Response Side, M = Match.

Formula	df	AIC	BIC	Log Likelihood	χ^2	Df	*p*
RT ~ 1 + L + (1|Item) + (1|Subject)	**5**	**49,979**	**50,010**	**−24,985**			
RT ~ VF + L + (1|Item) + (1|Subject)	**6**	**49,974**	**50,012**	**−24,981**	**6.92**	**1**	**0.009**
RT ~ RS + L + (1|Item) + (1|Subject)	6	49,980	50,017	−24,984	0.96	1	0.326
RT ~ M + L + (1|Item) + (1|Subject)	6	49,977	50,014	−24,982	4.27	1	0.039
RT ~ VF + RS + L + (1|Item) + (1|Subject)	7	49,975	50,019	−24,981	0.97	1	0.326
RT ~ VF + M + L + (1|Item) + (1|Subject)	**7**	**44,972**	**50,015**	**−24,979**	**4.27**	**1**	**0.039**
RT ~ VF + M + RS + L + (1|Item) + (1|Subject)	8	49,973	50,023	−24,979	0.96	1	0.328
RT ~ VFxM + L + (1|Item) + (1|Subject)	8	49,974	50,023	−24,979	0.20	1	0.655
RT ~ MxRS + VF + L + (1|Item) + (1|Subject)	9	49,975	50,031	−24,979	1.07	2	0.586
RT ~ VFxRS + M + L + (1|Item) + (1|Subject)	**9**	**49,968**	**50,024**	**−24,975**	**7.97**	**2**	**0.019**
RT ~ VFxRS + VFxM + L + (1|Item) + (1|Subject)	10	49,970	50,032	−24,975	0.19	1	0.667
RT ~ VF × RS + RS × M + L + (1|Item) + (1|Subject)	10	49,970	50,032	−24,975	0.11	1	0.736
RT ~ VF × RS × M + L + (1|Item) + (1|Subject)	12	49,973	50,048	−24,975	0.63	3	0.890

because of the lower AIC, the VF model was then compared to a model with an added fixed effect for match, and another model with an added fixed effect for response side, of which the model with fixed effects for VF and match proved to be the best fit ($\chi^2(1) = 4.27$, $p = .039$ and $\chi^2(1) = 0.97$, $p = .326$ respectively). Neither adding all three factors as fixed effects nor adding an interaction between VF and match yielded a significantly better fit compared to the model with main effects for VF and match ($\chi^2(1) = 0.96$, $p = .328$ and $\chi^2(1) = 0.20$, $p = .655$ respectively). Using a model with an interaction between match and response side and an added main effect of match also did not improve the fit ($\chi^2(2) = 1.07$, $p = .586$). However, a model with an interaction between VF and response side and an added main effect for match did improve the fit significantly ($\chi^2(2) = 7.97$, $p = .019$). This model also proved to be the best fit when compared to models where the main effect of match was replaced with a two-way interaction between VF and match ($\chi^2(1) = 0.19$, $p = .667$) or a two-way interaction between response side and match ($\chi^2(1) = 0.11$, $p = .736$), as well as the full model with a three-way interaction between VF, response side, and match ($\chi^2(3) = 0.63$, $p = .890$). To conclude, the data were described best by a model with an interaction between VF and response side and a main effect of match (VF: $\beta = 0.17$, $t = 0.02$, response side: $\beta = 47.87$, $t = 1.51$, match: $\beta = 13.71$, $t = 2.10$, VF x response side: $\beta = -34.66$, $t = -2.65$).[3] We used the same stepwise procedure to analyze accuracy rates. The null model with a fixed effect for length and random intercepts for participants and items (adding by-participant or by-item random slopes resulted in a failure to converge) proved to be the best fit for the data. The only model that came close to being superior ($p < .1$) was the model with an added VF by response side interaction ($\chi^2(3) = 6.50$, $p = .090$).

Because of the differences in word length between the German word pairs used in this experiment and the English word pairs used in Zwaan and Yaxley (2003b), we decided to conduct a post-hoc analysis excluding all trials in which one of the words of the word pair consisted of more than ten letters, since the maximum word length in the original study was ten letters. This led to the exclusion of 11.4% of the trials, and the resulting mean word length was 6.04 letters (compared to 5.75 in the original study). If the failure to find an interaction between match and visual field was due to the German words used in our replication attempt being longer, we should find the interaction of match

---

[3] The complete summary of the full model is: Intercept: $\beta = 674.24$, $t = 14.62$, VF: $\beta = 6.70$, $t = 0.52$, match: $\beta = 22.56$, $t = 1.72$, response side: $\beta = 53.86$, $t = 1.63$, length: $\beta = 10.56$, $t = 3.48$, VF x match: $\beta = -13.20$, $t = -0.71$,VF x response side: $\beta = -34.66$, $t = -2.65$, match x response: $\beta = -11.99$, $t = -0.65$, VF × match × response side: $\beta = 15.08$, $t = 0.58$.

and visual field when looking only at shorter words. Using the same stepwise procedure as was used in the main analysis, the best fit to the data was the model with main effects of visual field and match (VF: $\beta = -13.82$, $t = -1.99$, match: $\beta = 22.55$, $t = 3.25$, length: $\beta = 16.25$, $t = 4.16$). Thus, in contrast to the main analysis, we here did not observe an interaction of visual field and response side but rather a main effect of visual field. Importantly, however, as in the main analysis we observed a main effect of match but no interaction of match and visual field. Thus, word length does not seem to explain why no match by visual field interaction was observed in our main analysis. One might argue that the mean word length in the current analysis was still higher for the German words even when adjusting the range of the word length. We therefore repeated the analysis with trials in which words were at most nine letters long. This led to the exclusion of 22.3% of trials compared to using all word pairs, while the resulting mean word length was 5.75, the same as in Zwaan and Yaxley (2003b). The model with only a main effect of match proved to be the best fit to the data (match: $\beta = 22.47$, $t = 3.03$, length: $\beta = 17.59$, $t = 3.31$). Thus, in contrast to the main analysis, we did not observe an interaction of visual field and response side for these shorter words. Importantly however, replicating the results of the main analysis, we again observed a main effect of match that was not qualified by a match by visual field interaction. Thus, we feel quite safe in concluding that word length does not explain the differences between our replication study and the original study.

Another important difference between German and English is the fact that German nouns possess one of three genders (male, female or neutral). Although no articles were presented in front of the nouns, it could still be the case that differences in gender between the two words of the word pairs were perceived as a mismatch. To test this, wetested whether a match or mismatch of gender (both nouns possess the same gender or not respectively) would have any influence on the results. Adding this factor to the best model with an interaction between response side and visual field and an added main effect of match did not further improve the fit ($\chi^2(1) = 0.01$, $p = .933$).

4. Discussion

Our study was based on an experiment by Zwaan and Yaxley (2003b), who presented their participants with word pairs that were displayed in a vertical arrangement that either matched or mismatched the canonical spatial relationship between the referents of the words (e.g., “roof” above “basement”, or the other way around, respectively). Zwaan and Yaxley observed an interaction between the match effect and the visual field to which the word pair was presented. We used the same stimuli and methods as Zwaan and Yaxley (2003b) but added response side as an additional between-participants factor, in order to de-confound VF and contra- versus ipsi-lateral response side. Doing so, we no longer found a VF by match interaction, but instead only found a VF by response side interaction and a main effect of match.

We were not able to replicate the hemispheric asymmetry found by Zwaan and Yaxley (2003b), despite sufficient power. We had a sample size that was 2.5 times the size used in the original study, as suggested by Simonsohn (2015). Our results thus stand in conflict with the results of Zwaan and Yaxley regarding the VF by match interaction they observed: Instead of being confined to the RH, a match effect was found independent of VF and therefore independent of hemisphere. We also did not observe a three-way interaction of VF, response side, and match, which speaks against the above-mentioned possibility that match effects are only observed for stimuli that are presented contra-lateral to the required response and thus are typically associated with slower response times. To account for differences in word length between the English word pairs used by Zwaan and Yaxley (2003b) and the German word pairs used in our replication attempt (mean of 5.73 letters vs 6.34 letters respectively), we conducted post-hoc analyses with subsets of our data excluding trials that contained words with more than 10 letters, since that was the maximum word length in the original study, or 9 letters to bring the mean word length down to 5.75 letters to match the mean word length of the original study. We found a main effect of match and a main effect of presentation side when looking at the first subset and only a main effect of match when looking at the second subset but still no VF by match interaction. The stronger influence of match effect when looking at trials with shorter words indicates that participants were better able to identify and process the word pairs when they were shorter, but there did not seem to be processing differences between the hemispheres. Differences in gender of the German nouns also had no influence on performance.

How can the differences between the results in the current study and the original study by Zwaan and Yaxley (2003b) be explained then? One might argue that there was an attentional bias to the RVF in the experiment conducted by Zwaan and Yaxley (2003b): Both the response mode, with positive responses being assigned to the right hand, as well as the experimental setup could have biased the participant's attention to the RVF (Mondor & Bryden, 1992): Because of our left-to-right reading habit, we typically scan the LVF first when stimuli are being presented bilaterally. However, when stimuli are presented unilaterally and a central fixation point is shown before stimulus presentation, the attentional scanning starts at the fixation point and goes to the RVF first. This could have led to a RVF advantage in Zwaan and Yaxley (2003b), since the first letters of the word pairs, which are more important for word recognition compared to the last few letters (Balota & Rayner, 1991; Brysbaert, Vitu, & Schroyens, 1996), were attended to faster when appearing in the RVF compared to the LVF. Indeed, the match effect observed in the Zwaan and Yaxley study was actually due to a disadvantage in the mismatch condition rather than an advantage in the match condition. Only mismatching word pairs presented to the LVF were slower than the other three types of responses. Matching word pairs may have had a processing advantage compared to mismatching word pairs, supposedly independent of hemisphere, but since words presented to the RVF already had an attentional advantage, they did not further profit from the additional match effect leading to the specific result pattern observed by Zwaan and Yaxley (2003b). This attentional bias to the RVF could have been weakened in our experiment, due to counterbalancing of the response sides. However, it should be noted that in our experiment, we did not observe a three-way interaction of match, VF, and response side, which seems to speak against this possibility, considering that we manipulated response side between participants. In other words, if this explanation of the differences in results was correct, we should have observed a stronger match effect for word-pairs presented to the LVF (RH) for participants responding yes with their left hand than for participants responding yes with their right hand, which is not borne out by the data. Furthermore, it should be noted that when presenting word pairs in the center of the screen and thus having maximal attention, Zwaan and Yaxley (2003b) still found a match effect. This result also speaks against the idea that the differences in results obtained in our study and the study by Zwaan and Yaxley (2003b) can be explained in terms of a bias towards the right VF in the original study.

Interestingly, Zwaan and Yaxley (2004) also conducted an additional DVF study looking at the activation of shape information. Again using a semantic-relatedness task with only right-handed yes-responses, a word was presented centrally acting as a fixation point and after a while another word appeared either to the left or right of the already present word. In experimental trials, the two words were always unrelated (requiring a no-response) but either had the same or a different shape. Response latencies were higher when the words had the same shape (*“pie”* and *“tire”* for example) than when they had different shapes (*”pie”* and *“cheek”* for example), but this effect was only observed when the second word was presented to the RVF. In contrast to Zwaan and Yaxley (2003b), the match effect was thus confined to the LH instead of the RH in this experiment. However, considering that in this experiment experimental trials required left-handed no-responses, the match effect was again observed only when stimuli were presented contralateral to the required response. These results are consistent with the results of Zwaan and Yaxley (2003b) in observing match effects for stimuli that are presented contralateral to the required response. It is unclear why we did not observe such a pattern in our study but rather saw a main effect of match that was independent of VF. At the very least, however, the results of our current study and its implications in the light of previous experimental evidence demonstrate the importance of manipulating the response side in DVF studies.

There is still one problematic aspect that our replication study shares with the original study conducted by Zwaan and Yaxley (2003b): It is possible in principle that the stimuli were actually not presented completely unilaterally, since the middle of each word was presented in a distance of 2.67° visual angle to the left or right of a centrally presented fixation cross. In the original study, this led to the rightmost word of LVF items and the leftmost word of RVF items being only 1.49° visual angle away from the fixation point on average. Since there is supposedly an overlap between both visual fields in the foveal regions of our eyes, somewhere between 0.5° (Wyatt, 1978) and 3° (Bunt, Minckler, & Johanson, 1977). Bourne (2006) recommended that stimuli be shown at a distance of at least 2.5° to 3° visual angle from the fixation point when using a DVF paradigm. The short distance from the fixation point could have meant that at least some of the word pairs containing long words were projected to this foveal region, and therefore not presented unilaterally. Further DVF experiments are necessary to identify possible underlying factors that could have been responsible for the different results, controlling for the possibly insufficient distance from the center of the screen used in both the original study and our replication as well as including a mask after presentation of the word pairs to further control presentation duration of the word pairs.

Finding a Simon effect in our experiment underlines the importance of controlling for response side in DVF experiments. The absence of a VF by match interaction implies that there are no large differences between the hemispheres regarding the preference for processing of word pairs that match rather than mismatch the canonical spatial relationship between their referents. This casts doubt on the fine versus coarse semantic memory assumption proposed by Beeman (1998). Coney (2002) raised similar doubts about this distinction. He found no differences between the hemispheres in an associative priming paradigm, controlling the strength of the associative relationship. Stronger priming effects for weak associative strength should have been found in the RH, whereas for the LH the opposite should have been observed, according to the fine vs coarse distinction.

5. Conclusion

We attempted to replicate the VF by match interaction observed by Zwaan and Yaxley (2003b) using a DVF paradigm, where participants saw word pairs in a spatial arrangement either matching (*"nose"* written above *"mustache"*) or mismatching (*"mustache"* written above *"nose"*) the canonical spatial relationship between the referents of these words. In contrast to the original study, we counterbalanced response side between participants. No VF by match interaction could be observed; instead only a Simon effect (VF by response side interaction) and a main effect of match were found. This raises doubts with respect to the assumption that coarse semantic knowledge, including spatial relations, is confined to the RH. Future studies controlling for response side, word length, and attention while ensuring completely unilateral presentation are needed to investigate lateralization during language processing.

Declarations of Competing Interest

None.

Acknowledgements

This study was supported by the German Research Foundation (DFG) as part of the Collaborative Research Center "SFB 833"; Project B4 of Barbara Kaup and Hartmut Leuthold.

References

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.

Balota, D. A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables. In D. Besner, & G. W. Humphreys (Eds.). *Basic processes in reading: Visual word recognition* (pp. 198–232). Mahwah, NJ: Erlbaum.

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.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. Right hemisphere language comprehension: Perspectives from cognitive neuroscience. In M. Beeman, & C. Chiarrello (Eds.). *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255–284). Mahwah, NJ: Erlbaum.

Beeman, M. J., & Bowden, E. M. (2000). The right hemisphere maintains solution-related activation for yet-to-be-solved problems. *Memory & Cognition, 28*, 1231–1241.

Bourne, V. J. (2006). The divided visual field paradigm: Methodological considerations. *Laterality, 11*, 373–393.

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology, 10*, 385–395.

Bunt, A. H., Minckler, D. S., & Johanson, G. W. (1977). Demonstration of bilateral projection of the central retina of the monkey with horseradish peroxidase neuronography. *Journal of Comparative Neurology, 171*, 619–630.

Coney, J. (2002). The effect of associative strength on priming in the cerebral hemispheres. *Brain and Cognition, 50*, 234–241.

Dudschig, C., & Kaup, B. (2017). Is it all task-specific? The role of binary responses, verbal mediation, and saliency for eliciting language-space associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(2), 259–270.

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). Do task-irrelevant direction-associated motion verbs affect action planning? Evidence from a Stroop paradigm. *Memory & Cognition, 40*, 1081–1094.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science, 19*, 93–97.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*, 825–850.

Fodor, J. A. (1975). *The language of thought. Vol. 5*. Cambridge, MA: Harvard University Press.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: Automatic activation of location information during word processing. *Psychonomic Bulletin & Review, 18*, 1180–1188.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Langacker, R. W. (1998). On subjectification and grammaticization. In J.-P. Koenig (Ed.). *Discourse and cognition: Bridging the gap* (pp. 71–89). Stanford, CA: CSLI Publications.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315.

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science, 18*, 1007–1013.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48*, 788–804.

Mondor, T. A., & Bryden, M. P. (1992). On the relation between visual spatial attention and visual field asymmetries. *The Quarterly Journal of Experimental Psychology, 44*, 529–555.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28*, 292–312.

Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: On the relevance of events for perceptual simulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 579–590.

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences, 17*, 458–470.

Pulvermüller, F. (2018). Neural reuse of action perception circuits for language, concepts and communication. *Progress in Neurobiology, 160*, 1–44.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*, 351–360.

Pylyshyn, Z. W. (1986). *Meaning and cognitive structure: Issues in the computational theory of mind.* Norwood, NJ: Ablex.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*, 300–304.

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26*, 559–569.

Wyatt, H. J. (1978). Nasotemporal overlap and visual field sparing. *Investigative Ophthalmology and Visual Science, 17*, 1128–1130.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher, & R. Zwaan (Eds.). *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 224–245). Cambridge, UK: Cambridge University Press.

Zwaan, R. A., & Yaxley, R. H. (2003a). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10*, 954–958.

Zwaan, R. A., & Yaxley, R. H. (2003b). Hemispheric differences in semantic-relatedness judgments. *Cognition, 87*, B79–B86.

Zwaan, R. A., & Yaxley, R. H. (2004). Lateralization of object-shape information in semantic processing. *Cognition, 94*, B35–B43.