UiT
THE ARCTIC UNIVERSITY OF NORWAY

Faculty of health sciences, department of psychology

Words lead to action

Do incompatible movements interfere with verbal stimulus-response learning?

—

Are Skjoldli

Master's thesis in psychology - May 2018

Words lead to action: Do incompatible movements interfere with verbal stimulus-response learning?

Written by Are Skjoldli

PSY-3900 Master thesis

Word count: 7474

Spring 2018

UIT The Arctic University of Norway

Thesis advisor: Torsten Martiny-Huenger

Preface

I started studying psychology because I was interested in human interactions; what is causing us to behave as we do; and how we can work to better ourselves. In essence, to develop an understanding of myself, and of human behaviour in general. I was naturally interested in social psychology, in addition to cognitive sciences to develop further understanding of behaviour.

When I was looking for a project for my bachelor thesis, I was initially interested in the concept of free will, or rather how we can avoid falling victim to habitual tendencies. I found a project researching implementation intentions under the supervision of Torsten Martiny-Huenger. In short, it is a form of self-regulation method in which you make detailed action-plans (or if-then plans) aimed to achieve your goals. When you make such an action plan, you essentially prepare an action for a specific situation by means of mental simulation or verbal instruction. My previous thesis work was the main inspiration for my current thesis, since underlaying mechanisms in relation to language and actions is important to understand with regards to the theoretical framework of implementation intentions.

In my current thesis, I have found relevant research and written the entire thesis myself. I have also collected empirical data and performed the necessary statistical analysis'. Designing the experiment itself was more of a joint effort, in which I and Torsten Martiny-Huenger discussed different ways of how we could test a specific cognitive mechanism in a behavioural design. I have to give a special thanks to Torsten for writing the computer code necessary for running the experiment.

I wish to again thank Torsten in general for excellent advisement on everything related to the master thesis. I would also like to thank UIT The Arctic University of Norway for offering a master's program in which you are largely free do design your own project.

Student's signature

Thesis advisor's signature

Abstract

Embodied cognition research has shown that bodily activity could affect cognitive processes in multiple ways. It has for instance been theorized that when we process language related to physical actions such as the word “push”, we automatically use the motor-areas of the brain to simulate our understanding of this concept, much like actually performing a pushing action. In this study, I sought to test this causal link, and to further demonstrate that it is possible to interfere with the motoric processing of action-specific verbal cues by simultaneously executing physical actions that is the exact opposite of the actions directed by the verbal cues. Using a relational priming paradigm, 46 student participants was presented with visual cues, as well as verbal cues. According to embodied cognition theory, the individual verbal cues should direct the correct physical response to individual visual cues. Meanwhile, the participants were also performing a hand movement specific to each visual stimulus that was either compatible or incompatible with the presented action-specific verbal cues. I hypothesized that that I would find the proposed causal link between action-specific words and visual stimuli. Furthermore, I hypothesized that the item-specific probe-task would produce faster response times when the verbal cues and physical actions were compatible in the prime-task, and an interference effect when they were incompatible. The results suggest that a causal stimulus-response link between presented action-specific words and presented visual stimuli was formed in the learning task. However, I was unable to produce an interference effect, and was thus unable to determine a mechanism that could explain how this link was formed.

Abstrakt (Norwegian)

Embodied cognition forskning har vist at kroppslig aktivitet kan påvirke kognitive prosesser på forskjellige måter. Det har blitt foreslått at når vi prosesserer språk relatert til fysiske handlinger (som bl.a. ordet «dytt»), at vi automatisk bruker motor-områdene i hjernen for å simulere vår forståelse av konseptet. Denne motoriske aktiveringen skal ligne på den motoriske aktiveringen som skjer i hjernen når vi utfører en samsvarende fysisk handling i praksis. I denne studien, forsøkte jeg å finne den foreslåtte kausale forbindelsen mellom ord og handlinger, og videre demonstrere at det går an å forstyrre den motoriske prosesseringen av handlings-spesifikke verbale ord, ved å samtidig utføre fysiske handlinger som er motsatt av handlingene som blir indikert av de handlings-spesifikke ordene. Jeg anvendte et priming-oppsett, og presenterte 46 deltakere med stimuli par bestående av an visuell stimulus og en verbal stimulus. I følge «embodied cognition theory», kan det i teorien dannes assosiasjoner mellom handlinger indikert av individuelle handlings-spesifikke verbale ord og visuelle stimuli. Deltakerne utførte spesifikke hånd-bevegelser samtidig som hvert stimuli par ble presentert, som var enten samsvarende eller ikke-samsvarende med de handlings-spesifikke verbale ordene. Jeg dannet en hypotese om at jeg ville finne den foreslåtte kausale forbindelsen mellom handlings-spesifikke ord og visuelle stimuli. Jeg dannet også en hypotese om at jeg kom til å se raskere responstider i «probe»-oppgaven hvis utførte fysiske handlinger og handlings-spesifikke verbale ord samsvarte i «prime»-oppgaven, og i motsetning en interferens effekt når de ikke samsvarte. Resultatene foreslår at en kausal stimulus-respons link mellom visuelle stimuli og handling-spesifikke verbale ord ble formet i lærings-fasen av eksperimentet (priming). Jeg klarte i midlertid ikke å produsere en interferens effekt, og kan dermed ikke foreslå hvordan den kausale forbindelsen har oppstått.

Words lead to action: Do incompatible movements interfere with verbal stimulus-response learning?

The basic mechanisms of what role language plays in relation to actions remains an open question. The traditional views of cognition offer one explanation, but recent evidence suggest that a different approach might be in order. In this research area, increasing amounts of empirical evidence have been found in favour of the view proposed by embodied cognition theory, which has a core assumption that is in conflict with the one proposed by traditional cognitive theories (Barsalou, 2010).

The theoretical framework proposed by embodied cognition theory reveals new research options that is increasingly supported by empirical evidence. I want to use this framework to investigate the relation between language and action, specifically regarding how action related words (e.g. push) can influence physical actions in response to stimuli, also known as stimulus-response learning.

Since much research in favour of the embodied cognition view is done to disprove the mechanisms postulated by traditional theories, a short introduction of the traditional view is in order for a better understanding of the theoretical framework the current experiment is based on. It will be followed by a broader introduction of embodied cognition theory which explicate the cognitive mechanisms I seek to investigate. When the theoretical framework is established, the experiment will be introduced.

Traditional cognitive theories

The most dominant view of cognition is provided by traditional cognitive theories, which use a computer analogy to explain the mechanisms underlaying cognition. According to traditional cognitive theories, acquired sensory information (perceptual, motoric and introspective) is first processed in the modalities of the brain, then coded as representations (often as words or sentences) into a separate semantic memory system following syntactic rules. To clarify, brain modalities are areas of the brain that is thought to have one primary function, such as the visual cortex which process sensory information with visual properties. The semantic memory system effectively represents the collected world knowledge we have developed from experiences, and when we retrieve knowledge for practical use, we construct higher order cognition based on our semantic representations (Barsalou, 1999). When we for instance retrieve knowledge about a specific item category from memory (e.g. a car), we use our stored semantic representations to construct relevant conceptual knowledge for applied settings (Barsalou, 2005; Barsalou, Simmons, Barbey and Wilson., 2003).

The representations contained in this semantic memory system is referred to as "amodal" as they are syntactic in nature and thus different from modality-specific representations (e.g. vision). To clarify, traditional views differentiate between amodal representations and modality-specific ones: Modality-specific representations is visual and motoric in nature and separate from the amodal knowledge system. Even though an amodal representation can be a product of perception, it is still syntactic in nature and thus unrelated to modality-specific functions (e.g. vision). This distinction can better be explained if we imagine an image (e.g. a birthday party). Using visual imagery/representations, we can imagine a cake, some guests, and some balloons. Imagining this scenario is an example of visual representation, which is modality-specific to the visual brain systems. However, if we want to identify "items" in the image, we have to use our amodal symbols that represents our knowledge. This way we can know that the cake is made of cream, and that one of the guests is called John (Smith & Kosslyn, 2014, p.165-180). In this way, amodal symbols are separated from the modality-specific representations (Barsalou, 1999).

If one were to investigate the relation between language and action using the models of traditional cognitive theories, our knowledge of action-words (e.g. kick or push) would be derived from stored semantic information. When we hear an action-word, information regarding this action is constructed from semantic memory to be applied in a contextual setting (Barsalou et al., 2003). Since this knowledge is theorized to be located in a specific brain area, this is presumably where we would see brain activity as a result of hearing action-words if we used brain imaging techniques. This means that we would not see any activity in the motoric or visual areas as these systems will be unrelated to retrieval of amodal knowledge (Barsalou, 1999).

Recent evidence does however suggest that the traditional models may be limited, which in turn puts its core assumption of a semantic memory system into question (Barsalou et al., 2003). The notion of a semantic memory system has been scrutinized for different reasons (e.g. Fodor, 2000). For one, that its supposed location in the brain is still debated (Barsalou, 1999). It also means that the theories are problematic in that an abstract theory regarding a semantic memory system is difficult to disprove (see Stanfield & Zwaan, 2001 for a similar argument).

Embodied Cognition

More recently, a theory was devised that aimed to give an alternate explanation of the basics underlaying mechanisms of cognition. This theory is called embodied cognition theory and offers a radically different view to that of traditional cognitive theories. According to this view, all the modalities of the brain work together to process and make sense of our sensory experience, instead of knowledge being stored in a syntax based semantic memory system. It is theorized that experiences instead are stored in association areas of the brain that can later be retrieved as partial reconstructions of previous experiences. This “knowledge retrieval” happens in a process referred to as mental simulation (Barsalou, 2005, 2008).

The mental simulations that effectively represent knowledge can be both conscious and unconscious. When we consciously simulate a scenario, we draw on previously stored sensory experiences to construct higher order cognition (e.g. though). Instead of retrieving amodal symbols from a memory system, brain activation from experiences is (simply put) re-enacted. For instance, let’s say you have reached the top of a mountain once (e.g. Mount Everest). When you mentally simulate what this experience was like, you partially re-create the sensory experiences you had at the time (Barsaou, 2005).

However, most of the time the mental simulations happen unconsciously to make sense of sensory information. For instance, if you hear an action-specific word (e.g. push) a mental simulation is theorized to happen in the motor areas of the brain to convey practical meaning to this word (e.g. to push a drawer shut). The mental simulation does not have to be specific to one modality. If you similarly were to hear the word “eagle”, a mental simulation will happen in the visual areas of the brain (if you have a visual referent that is), but there could also be additional activations in the other modalities based on your previous experiences. If you for instance know what sound an eagle makes it could trigger an auditory simulation, and if you have petted an eagle in the past it could trigger a somatosensory activation. Depending on your previous experience, all the modalities work together both consciously and unconsciously to re-enact sensorimotor representations through mental simulation. This is what makes out your conceptual knowledge according to the embodied view (Barsalou, 2005).

It is important to note that mental simulations are not theorized to be exact re-creations of experiences, but rather partial re-enactments. Mental simulations should be more accurate if you have more experiences to draw from. Let’s say you wanted to mentally simulate the process of buying an apple the next time you find yourself in the school cafeteria. To achieve accuracy in this simulation, it is essential to have been in the cafeteria before, and it is even better to have been there every day for the past year. You should thus be able to simulate a sensory experience that is very close to reality. Conversely, it is entirely possible to mentally simulate a scenario you have never experiences before. In which case, you use the contextually relevant experiences available to best achieve accuracy (Barsaou, 2005).

The main difference between embodied views and traditional views is how knowledge is represented in the brain (Barsalou, 2010). In summary, that traditional views postulate that knowledge is stored as language based (following syntactic rules) representations called amodal symbols which is retrieved to support the higher cognitive functions such as: Thought, language, memory and knowledge. Meanwhile, embodied views postulate that experiences are stored in association areas of the brain and are activated through mental simulation to convey meaning to sensory information. Higher cognitive functions are supported as a result of conscious mental simulations that is grounded in experiences. In the embodied view, knowledge is thus grounded in mental simulations (Barsalou, 1999, 2005, 2008).

Language in relation to motoric activation

Because my current research is on the language-action relation, I will now focus more on the motor-control aspect of embodied cognition theory. Specifically, I want to investigate the underlaying mechanisms of verbal language in relation to actions initiation. This research question has been investigated directly or indirectly in the past. By indirectly, I mean that the researcher(s) did not necessarily use embodied cognition theory as an underlaying assumption, but that the research is relevant for the research question nonetheless.

The core assumption of embodied views with regards to how knowledge is represented in the brain is grounded in simulations based on experience. This means that unconscious mental simulations related to actions is a re-enactment of physical actions you have performed in the past. In other words, when we for instance are presented with words related to actions (e.g. "push") it should trigger a mental simulation of actions performed in the past that is contextually related to the present situation. Using the same example, hearing the word "push" should in turn activate motor areas of the brain that is related to arm movements as this is where motoric activity occurs when you perform a pushing action in practice (Barsalou, 2005). Using the traditional views of cognition, we would not have been able to justify this assumption as retrieval of syntactic representations (amodal symbols) should be unrelated to motoric brain processes (Barsalou, 1999).

We will now look at empirical evidence which suggest that simulation mechanisms convey meaning to language and provides insight into how this process occurs.

Behavioural evidence. Glenberg and Kaschak (2002) investigated the interaction between sentence comprehension and physical action responses. As an underlaying assumption of the experiment they used "the indexical hypothesis" which proposes that linguistic meaning is a direct result of possible actions towards the environment, rather than the arbitrary abstract symbols proposed by traditional cognitive theories (Glenberg & Robertson, 1999, 2000). The notion that meaning is derived from action, is aligned with the explanation provided by embodied cognition theory to explain the workings of language in the brain (Barsalou, 2008).

In three experiments, Glenberg and Kaschak (2002) presented participants with sentences indicating a directional action, or sentences that had no practical meaning. For instance, participants would be presented with the sentences: "Open the drawer" and "close the drawer". The former indicate that you have to pull your arm towards yourself to open the drawer, and the latter, that you have to extend your arm to close the drawer. Participants would also be presented with sentences such as: "Courtney handed you the notebook", "you handed Courtney the notebook", "Liz told you the story" and "you told Liz the story", of which the former pair implies a directional transfer between you and another party, and the later indicates an abstract directional concept. The participants were also presented with directionless and meaningless sentences such as: "Boil the air".

The Participants were tasked with answering whether the sentences made sense by pressing two buttons on a custom-made keyboard with one response-button that had to be reached for (arm extension), and one that was close to your body (arm flexion). Thus, the two possible responses were designed to relate to the actions indicated by the sentences. The far-button was assigned the label "yes" (the sentence makes sense) and the near-button "no" (the sentence does not make sense). Halfway through the experiment, the buttons were reversed (Glenberg & Kaschak, 2002).

The results showed that the direction indicated by the presented sentences facilitated faster response times in the indicated direction. For example, if the sentence "close the drawer" was presented, participants responded faster if the button for yes was at the far side of the keyboard, and comparably slower if the button for yes was at the near side of the keyboard. This effect was seen regardless whether the sentences included concrete or abstract directions (Glenberg & Kaschak, 2002).

These results show that sentence understanding to some degree is grounded in action as proposed by embodied cognition theory (Barsalou, 2008). The fact that opposite responses seems to suggest an interference effect (slower responses) indicate that additional cognitive resources may be required to convey meaning and thus be able to make the right response (Glenberg & Kaschak, 2002).

The results of this study seem to suggest that understanding derived from a sentence will facilitate relevant motoric responses in line with that understanding. This is critical observation, since my experiment build on this assumption. In my experiment, I incorporate a stimulus-response paradigm that require a directional motoric understanding of action words. Although Glenberg and Kaschak (2002) used directional sentences to contextually imply meaning, it is still highly relevant as the meaning of the action words in sentences is an important component that often decides the directional property (e.g. give/take, push/pull).

Neuroimaging evidence. As simulation is the core assumption of how knowledge is represented in the brain in the embodied view, brain-imaging techniques offers a way to observe this interaction. If action-words are simulated in motor areas to convey meaning, it should show as motoric activation in the relevant brain areas. In other words, since performing a pushing action in practice activates brain areas related to arm-movements and simulation of pushing actions is grounded in previous experience, it should activate the same brain areas when hearing the action word as those that are otherwise activated when actually performing a pushing action with your arm(s) (Barsalou, 2005).

A neuroimaging experiment was done by Hauk and Pulvermüller (2004) to investigate the relation between action-oriented words (e.g. kick, lick, write) and cortical brain activation. The authors hypothesized that action-specific words related to leg-actions, face-actions and arm-actions respectively, would activate cortical brain areas corresponding to each word category. In other words, brain areas that would otherwise be activated when performing these actions in practise should also be activated in a similar manner by reading action-specific words.

The researchers used electroencephalography (EEG) which measures cortical electrical activity through electrodes placed on the participants' scalps (Hauk & Pulvermüller, 2004). Based on data collected in a previous study (Pulvermüller, Shtyrov and Ilmoniemi, 2003), the authors expected that a relevant peak in cortical activity would appear in frontal areas of the brain approximately 220ms after stimulus onset. In other words, approximately 220ms after a participant is presented with an action related word, motoric brain areas will activate as a response to the presented stimulus. The participants observed word-presentations on a computer screen without responding to the stimuli in any way and were encouraged to remain still throughout the experiment trials to reduce motoric brain activity (Hauk & Pulvermüller, 2004).

By analysing the EEG readings, the researchers were able to determine the mean brain activation for each participant between 210 and 230ms after stimulus onset. The results showed that there were in fact motoric activation differentiating between each word category as the authors predicted. This means that arm-related words activated areas of the brain related to arm-movements, face-related words activated areas related to face-movements, and leg-related words activated areas related to leg-movements (Hauk & Pulvermüller, 2004). These results support the model proposed by embodied cognition theory, which postulate a connection between the understanding of word-meaning and corresponding motoric brain activation (Barsalou, et al., 2003; Barsalou, 2005).

A following neuroimaging experiment done by Pulvermüller, Hauk, Nikulin and Ilmoniemi (2005) used transcranial magnetic stimulation (TMS) to test whether stimulation in the relevant areas of the motor-cortex would facilitate faster lexical responses to action-specific words. The stimulation was applied to the left language-dominant hemisphere of right handed participants for two reasons: 1) Because motoric processing of physical actions happens in the opposite brain hemisphere (contra lateral processing), which means that performing a physical action with your right foot (e.g. kick a ball) will activate the relevant area of the motor cortex in the left hemisphere and vice versa; 2) the left hemisphere of the brain is known to be essential for language processing.

The results showed that TMS in the motor areas related to arm-actions facilitated faster response times towards words related to arm-actions (e.g. “pick” and “grasp”) compared to words related to leg-actions, (e.g. “kick” and “step”) and neutral words. The reversed effect was seen when TMS was applied to leg areas. The TMS induced response times was also faster than the control conditions in which the stimulation was applied to the right hemisphere or the stimulation was faked on the left hemisphere.

These results suggest that action-related language is causally linked to motoric brain-activation (Hauk & Pulvermüller, 2004), and that this activation will facilitate relevant physical action initiation (Pulvermüller et al., 2005). This suggests that language understanding to some extent is grounded in motoric brain activation as embodied cognition theory suggests (Barsalou, 2008). With regards to the research done by Glenberg and Kaschak (2002), we also know that understanding of actions in different contexts (e.g. relative direction) could be important to facilitate an effective response.

Hauk and Pulvermüller (2004) showed that hearing an action specific-word (e.g. kick, push) facilitates motoric activation in the same areas of the brain that are otherwise active while performing said action in practise. In the following experiment they also showed that stimulation of motoric areas in the brain facilitated faster reactions in response to presented relevant action-related words (Pulvermüller et al., 2005).

The language-motor relation shown by Pulvermüller et al. (2005) is very important for the stimulus-response learning that will be introduced in the next section. In the traditional view, stimulus-response links are created by actually performing an action in a specific situation. However, these experiments and the model proposed by embodied cognition theory seem to suggest that a similar link can also be formed by action-related language (Barsalou, 2008; Glenberg & Kaschak, 2002; Hauk & Pulvermüller, 2004; Pulvermüller et al., 2005).

It would be interesting to further investigate if hearing and reading action-oriented words will link a response to a specific stimulus. To put it in an applied setting: What if you could train an action in response to a specific stimulus simply by reading an instruction manual or a self-written letter? The experiments reviewed thus far used visually presented words that the participants had to read. Will we see a similar effect if participants are hearing words (audio presentation) instead of reading them? This means, the stimulus would not be associated with the verbal aspects of the response (response word) but actually with motor related aspects representing the verbal content. The next experiment investigates these questions.

Stimulus-response associations

If we want to look closer at the effects words have on action, we can look to research on stimulus-response associations According to traditional views, association is first formed, then strengthened as a result of repeated actions made in response to a stimulus. This process can be illustrated if you can imagine the process of learning to drive a car. At first, you have to pay extra close attention to everything that happens around you while struggling to operate the vehicle. After a while though, you learn to recognise what you have to pay special attention to, and what you can ignore while driving. Driving itself gradually becomes more automatic and will in time be nearly effortless. This is a result of stimulus-response associations becoming strengthened and automated over time (Logan, 1990).

Recently, a study was done by Pfeuffer, Pfister, Moutsopoulou, Waszak and Kiesel (2017) that aimed to promote onset of stimulus-response associations in absence of action. In three experiments, the authors used a priming paradigm to make participants establish stimulus-response associations. In the priming trials, the participants established stimulus-response associations between pictures of objects, and a correct key press response (pressing a key on a computer keyboard). In the probe task, the participants had to press the correct key in response to the pictures presented in the previous priming trials. For instance, a participant could be presented with a picture of a car and told to categorize it as big or small. In the priming task they would categorize this picture by pressing “a right key” or “a left key”, and thus form an association between car-picture and the correct response. When the car-picture was presented again in a probe trial the participant should be able to categorize the picture quicker since the association between the picture and the response had been formed in a previous prime trial. However, the authors also designed a priming condition in which the participants observed the prime stimuli (car-picture), but the correct response was played over headphones. For instance: “Big”, “right key”, the former directing the categorization criteria and the latter directing the correct response. The probe task was not always a direct repetition as in the example above, in some trials the response buttons could be switched, the classification criteria changed, or both (e.g. “Mechanical”, “left key”). This was introduced to eliminate learning effects.

The results show that stimulus-response associations were found in both conditions: 1) when the priming block included physical responses, and 2) when the priming block included passive listening to audio instructions. This was unexpected with regards to stimulus-response association theory (Pfeuffer et al. 2017) but is in line with embodied cognition theory (Martiny-Huenger, Martiny, Park-Stamm, Pfeiffer and Gollwitzer, 2017). According to stimulus-response association theory, an association is formed by the co-occurrence of a stimulus and an action. Since the second condition did not include action execution it should not have formed an association, but the results suggest otherwise (Pfeuffer et al., 2017). Meanwhile, embodied cognition theory propose that action-specific words will activate relevant motoric areas of the brain depending upon subjective understanding of the context (Barsalou, 2008). As we have seen, Pulvermüller et al. (2005) showed that motoric TMS to the motor cortex facilitated faster response times in response to action-specific words, not unlike the priming effect used in this experiment. Glenberg & Kaschak (2002) also showed that the understanding of the context is relevant for facilitating actions. This can be related to this experiment since the primed responses included the phrases “right key” and “left key”, not unlike the directional sentences used in Glenberg & Kaschak (2002) to imply response direction.

Experiment introduction

In my experiment, I sought to further investigate the proposed link between language and action. The results from the experiment done by Pfeuffer et al. (2017) suggest that presenting visual stimuli together with verbally presented action-specific words may be sufficient in order to establish a stimulus-response link. I partially replicated the study done by Pfeuffer et al. (2017), but with some adjustments designed to test the underlaying assumptions of embodied cognition theory. In the original experiment, the researchers used key-pressing as the way for participants to respond to stimuli. This was changed in my experiment to put more emphasis on motoric responses and to further align my research with the studies reviewed thus far (Glenberg & Kaschak, 2002; Hauk & Pulvermüller, 2004; Pulvermüller et al., 2005). Instead of key-responses as a way to respond to stimuli, my experiment incorporated both a computer-mouse and a joystick.

I incorporated a similar paradigm to that which was used by Pfeuffer et al (2017). I designed one task to form associative links between visual stimuli and verbal cues (learning phase), and a follow-up task to test if associations had been formed (test phase). Unlike the experiment done by Pfeuffer et al. (2017), I also tested if it is possible to interfere with formed associations. I did this to further test the theoretical assumption that action-specific words trigger a motoric activation similar to that of performing an action (Barsalou, 2005).

In the current experiment, I designed a learning phase in which visual stimuli were presented alongside the action specific verbal cues “push” and “pull” in order to form stimulus-response links. The visual stimuli presented in the learning phase was later presented again in a following test phase, in which the participants were instructed to respond to them as quickly as they could. I hypothesized that the learning phase would facilitate faster response times in response to stimuli (as it would be the case with regards to stimulus-response associations) in the following test phase.

Furthermore, I tested if performing arm movements at the same time as the action-specific verbal cues were presented would affect the learning effect in any way. To achieve this, I chose to design the learning phase in a way that made the participants perform pushing or pulling movements with a computer-mouse while being focused on the presented stimuli. The computer program was designed so that the arm movements would be compatible with the verbal cues 50% of the time (e.g. hearing the word “push” and meanwhile performing a pushing action), and otherwise incompatible (e.g. hearing the word “push” and meanwhile performing a pulling action). I hypothesized that incompatible arm movements would interfere with the learning effect of the presented verbal cues. The suggested similarity between conceptual understanding of action-specific words and performing physical arm movements, would make it reasonable to assume that an interference effect might occur if these were to happen at the same time (Barsalou, 2005).

In the following test phase, the participants categorized the previously presented visual stimuli by pushing or pulling a joystick. I further hypothesized that the participants would respond faster to stimuli that had been paired with compatible arm movements in the previous learning phase (learning effect occurred), and slower to stimuli that had been paired with incompatible arm movements (interference with learning effect). I also included neutral stimuli that was never shown in the learning phase to act as a control condition.

According to the evidence reviewed thus far, an action-specific word should trigger a motoric activation in the brain and facilitate a relevant motoric response (Hauk & Pulvermüller, 2004; Pulvermüller et al., 2005). The reason for this is theorized to be that our understanding of action-words is grounded in specific motoric brain simulations (Barsalou, 2005). Since stimulus-response associations is a well-established concept (Pfeuffer et al. 2017), it could be possible to interfere with a verbally induced motoric learning effect by having a participant perform a physical action that are the exact opposite of the verbally presented action-word.

My hypothesis can be illustrated in three steps: 1) The word "push" activates areas of the brain related to arm-movements; 2) however, a participant might simultaneously perform a pulling action at the same time which also activates these brain areas; 3) since these activations are in conflict with each other, it could create an interference effect in which they cancel each other out, and thus a motoric priming effect is not formed.

I thus made the following predictions: 1) that action-specific words presented with visual stimuli in the learning phase, would form a link between each stimulus and the motoric response indicated by an accompanying action-word ("push" or "pull"). This link should in turn facilitate faster responses to the same visual stimuli presented later in the test phase as long as there is no physical interference present; 2) that incompatible arm-movement performed simultaneously as the action-words were presented in the learning phase, would interfere with the formation of a stimulus-response link and response times in the test phase would be comparably slower as a result.

Method

Participants and design

I recruited 46 participants (16 male and 30 female) from the university of Tromsø (Norway), with a mean age of 22.30 (SD = 2.86) years ranging from 18 to 31. All the participants signed a consent form to participate. The design followed a 3-factorial (stimulus type: compatible vs. incompatible vs. control) within-participant design with the joystick response times in the test phase as the dependent variable.

Hardware and software

All participants worked on the same stationary computer with a 3.30GHz Intel Core I5-4590 processor, 8GB RAM, and a windows 10 enterprise 64-bit operating system. The necessary external devices included: An optic mouse, a joystick and a speaker. I used the Psychopy software (Peirce, 2009) to program and run the whole experiment.

Procedure

When starting the experiment, the participants were informed that the purpose of the study was to measure memory after performing a task involving difficult multitasking. They were then informed that the experiment consisted of two main tasks, and that their performance in the learning phase would better their performance in the test phase. At the end of the experiment introduction, the participants were given a complete layout of the tasks in the experiment: 1) test phase practice block, 2) learning phase practice block, 3) learning phase, 4) test phase. They were also told that the experiment was structured in this way so that they would understand how the learning phase and the test phase were connected to each other.

Test phase practice block. In the test phase practice block, the participants were told to categorize images of objects based on their size (small or large) by using the joystick to push or pull in response to stimuli. They were given a reference point to determine the definition of the term large and small: "Would this item fit inside a shoebox"? Every time the participants responded correctly to the stimuli, the auditory verbal cue "push" or "pull" (depending on the correct response) would be presented over the speaker. There was no time-limit to complete the task, and a total of 10 practice trials were presented. If a participant moved the joystick in an undesirable manner (i.e. using he's/her wrist to move the joystick) the experimenter would comment on this after the practice task was completed.

Learning phase practice block. In the learning phase practice block, the participants were instructed to move the mouse cursor between the corners of computer screen depending on which corner was highlighted with a blue colour. When the mouse cursor was moved to the highlighted corner, a new corner would be highlighted, and they would repeat the process. Participants had to move the mouse cursor to 10 highlighted corners during this practice task. If a participant moved the mouse in an undesirable manner (i.e. not goal oriented movements or dragging the mouse with wrist movements instead of using wider arm movements), the experimenter would comment on this after the practice task was completed.

Learning phase. After the participants had read the instructions for the learning phase, the experimenter answered any questions before leaving the room.

In the learning phase, the participants were instructed to move the mouse-cursor to corners of the screen highlighted in blue and pay close attention to presented stimuli while performing the task. The participants would have to move the mouse to a few seemingly random corners of the screen before the first pair of visual and audio cues was presented. When a stimuli presentation occurred, it included an image of a large or small object appearing in the middle of the screen, as well as an auditory verbal cue that said either “push” or “pull”. The verbal cue indicated the correct response to the presented stimuli in the later test phase, however the mouse movement required to move the mouse cursor to the next highlighted corner was either compatible or incompatible with the action indicated by the verbal cue. The verbal cues and arm-movements were compatible 50% of the time, and otherwise incompatible 50% of the time. For example, in an incompatible trial the verbal cue “push” could be presented, but meanwhile the participant was performing a pulling action in order to move the mouse to the next highlighted corner. In all, 24 pairs of visual and audio cues with 3 randomized repetitions (72 presentations in total) were presented in the learning phase.

Participants had to continuously move the mouse cursor between the corners of the screen, and each trial consisted of 4 arm-movements. In each trial the stimuli (verbal cue and object image) were presented when moving the mouse from the second highlighted corner to the third. The visual cue were presented when the mouse cursor reached the second highlighted corner, and the verbal cue was presented when the cursor was moved one third of the distance (on the screen) towards the third highlighted corner. The movement direction required to move the cursor between the second and third corner was always vertical (push or pull movement). The audio cue was played once and lasted for approximately 300ms (either “push” or “pull”), and the visual cue lasted for more than 1000ms before disappearing from the screen. The length of the visual cue presentation was dependent on when the verbal cue was presented, which in turn was dependent on when a participant had moved the mouse one third of the way from the second corner towards the third corner. Each visual stimulus always disappeared 1000ms after the presentation of the verbal cue was initiated.

After a random set of trials, the participants were asked to remember the last picture presented and the verbal cue associated with it to make sure the participants did not simply ignore the presented cues.

The mouse-cursor speed (mouse sensitivity) were set on the lowest setting in windows. This way, the action-specific verbal cues would represent (or not) the performed arm-movements more than moving a computer-mouse in a casual setting (more wrist than arm movements).

Test phase. Like in the previous test phase practice block, the participants were told to categorize images that appeared at the screen as large or small by using the joystick to push or pull in response to stimuli. Unlike the previous practice block, they were also told to categorize the images as fast and as correct as possible. They were again given the reference point to determine the definition of the term large and small: “Would this item fit inside a shoebox”?

The participants were presented with the 24 images that had been presented earlier in the learning phase. The verbal cues presented alongside the pictures in the learning phase and the correct answer (using the joystick to either push or pull) were always compatible. The participants also had to categorize 12 additional neutral stimuli that were never presented in the learning phase. Each image was only presented once (no repetitions).

Before every trial, a fixation cross would appear on the screen for 500ms followed by a blank screen for 300ms indicating that a stimulus would appear soon. Then a visual stimulus was presented and last for 5000ms. If a correct response was made, the visual stimulus would disappear, and the process would start anew. If a wrong response was made it, would result in an error message telling the participant that they made a wrong response. Additionally, if no response was made for 5000ms it would also result in an error message telling the participant to respond faster. An error message lasted for >2300ms.

The joystick was fastened to the table, so it would not move around during the categorization trials. Furthermore, the experimenter helped positioning the joystick before the task, so every participant could use it comfortably, and with their dominant hand. The joystick had to be deflected by 80% of its total range in order for it to count as a response, and the next trial was only initiated after the joystick had been returned to its starting position.

End questionnaire

At the end of the experiment, the participants were presented with a short questionnaire intended to gather necessary demographical information. The questionnaire also included specific questions regarding the experiment itself such as: “Are you left or right handed” and “which hand did you use to move the mouse and joystick during the experiment”.

Debriefing

After the experiment, the participants were given elementary information regarding my study and what I was investigating, as well as the fundamental theoretical framework (e.g.

embodied cognition theory) to put the experiment into a comprehensible context. The participants were also asked to not share critical information about the study to other potential participants for the duration of the data collection process.

Data treatment and analysis

All trials in which the participants had responded wrong, or failed to respond faster than 5000ms, were excluded from further analysis (3.6%). Furthermore, response-times that was greater than 3 times the standard deviation of the individual mean response-times for each participant were also excluded from further analysis (1.9%).

To analyse the data and produce plots, I used a one-way repeated measures analysis of variance (ANOVA) to assess the relationship between response times in the three conditions (compatible vs incompatible vs control). Furthermore, I used a paired samples t-test to assess the differences between the three conditions.

Results

A one way repeated measures ANOVA showed a significant difference between the the three stimuli conditions (compatible vs incompatible vs control), $F(1, 45) = 17.58$, $p < .001$, $\eta^2 = .444$.

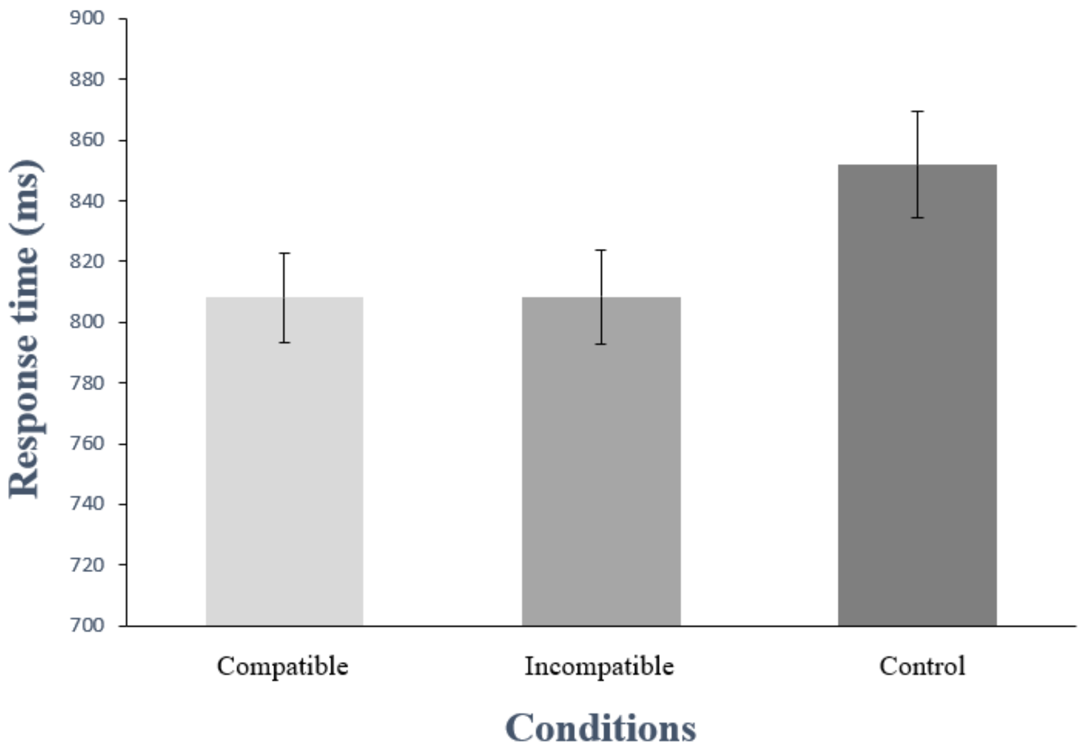

Figure 1. *Mean joystick response-times in milliseconds for the 3 stimuli conditions: Compatible, incompatible and control. Error bars show mean standard error for each condition.*

The result of the paired sample t-test showed that the action-specific verbal cues in the learning phase facilitated faster response times in the test phase of the experiment, thus confirming my first hypothesis. Responses to stimuli in which arm-movements were compatible in the learning phase M = 808.04ms SD = 99.45ms were significantly faster than control stimuli M = 851.77ms SD = 118.74ms, BCa 95% CI [-61.80, -25.67] $t(45) = -4.86$, $p < .001$ (two tailed). Responses to stimuli in which the arm-movements were incompatible in the learning phase M = 808.31ms SD = 104.47ms was also significantly faster than control stimuli M = 851.77ms SD = 118.74ms, BCa 95% CI [-61.28, -25.62] $t(45) = -4.91$, $p < .001$ (two tailed). There was however no difference between response times to stimuli that had been presented in the learning phase, despite my attempt to induce interfering arm-movements in 50% of stimuli presentations. In other words, I was unable to show an interference effect, which means that my second hypothesis was not confirmed.

Discussion

In the present research, I have used the theoretical framework proposed by embodied cognition theory to investigate a proposed stimulus-response link between action-specific language and action facilitation (Barsalou, 2008). I have also reviewed several studies from different fields of psychology as empirical support, and to put my assumptions in context (Glenberg & Kaschak, 2002; Hauk & Pulvermüller, 2004; Pfeuffer et al., 2017; Pulvermüller et al., 2005). In summary, this theoretical framework suggest that our understanding of words trigger a mental simulation based on previous experience. This means that hearing the action-word "push" should facilitate a motoric simulation similar to pushing actions you have performed in the past, as this is where your understanding of the word is grounded. Since this motoric simulation is grounded in experience, the brain activity should be similar to that of actually performing a physical action. Thus, repeated mental simulation in response to stimuli should in theory have similar effects to that of stimulus-response associations.

Associative learning hypotheses

I expected that a motoric link would be formed between presented action-specific words and presented visual stimuli in the learning phase of the experiment, similar to the language-motor link shown by Pfeuffer et al. (2017). Since the verbal cues that was paired with visual stimuli in the learning phase always indicated the right response for the later test phase, it suggests that the action-words facilitated faster response times in response to relevant stimuli later in the test phase (compatible: 808.04ms and incompatible: 808.31ms vs control: 851.77ms). The theoretical framework provided by embodied cognition theory explains the results as a stimulus-response link formed between stimuli and action-words (Barsalou, 2005), suggesting that my first hypothesis is correct. However, there is an alternative explanation that also needs to be addressed.

Another explanation is that the participants simply recognised the recurring images presented in the test phase and were thus able to categorize them faster than the control images. The visual stimuli in the learning phase were presented with three repetitions before being presented once in the test phase, in contrast the control stimuli were only presented once in the test phase. It is logical to assume that that seeing a new image (a control) for the first time would require a moment in order to properly identify the pictured object and categorize it accordingly, compared to a pictured object you have already seen with 3 repetitions (e.g. Breuer, Masson, Cohen & Lindsay, 2009).

The associative learning hypothesis I made appears to be supported by my findings. However, the fact that the participants were repeatedly exposed to the visual stimuli in the learning phase while the control stimuli were only shown once in the test phase, puts my control condition into question. In the test phase, the control stimuli could have taken longer to identify and categorize accordingly, hence the longer response times. This is an issue that could make the confirmed hypothesis less valid as the slower response times could have an alternate explanation. This could be addressed in future experiments by adding visual stimuli in the learning phase (or an equivalent learning task) that is not pared with a verbal cue. If response times to these stimuli are slower than those of a comparatively compatible condition, it would suggest that the associative learning hypothesis is correct.

Motoric interference hypothesis

It was unexpected that the arm-movements the participants performed in the learning phase did not appear to influence the response times in the test phase. That the co-occurrence of a stimulus followed by a physical action creates and strengthens a stimulus-response association is a well-established concept (Logan, 1990). However, we did not observe anything to suggest that the physical arm-movements in the learning phase had any effect on response times in the test phase.

The interference effect I attempted to induce should in theory have similar properties to that shown by Glenberg and Kaschak (2002). When the participants in their experiments were presented with sentences indicating a response direction, incompatible response movements seemed to interfere with the participants ability to respond efficiently. By explaining the results of their experiments within the theoretical framework of embodied cognition theory, it would suggest that the mental simulations related to understanding the presented sentences and their directional context were incompatible with the simulations involved in opposite responses directions. I expected to see a similar effect in my experiment, if the physical arm-movements in the learning phase had cancelled out or overwritten the stimulus-response associations formed between visual stimuli and verbal cues. However, incompatible arm-movements in the learning phase did not appear to have any interference effect on the formation of stimulus response links. In other words, my second hypothesis was not confirmed.

An alternative explanation for why an interference effect was not observed, could be that the physical movement in the learning phase was not precise enough. A mouse tracking program was incorporated into the software used to run the experiment. It was designed to measure the direction of the mouse movements in the learning phase. I expected the directions in the mouse-tracking data to be straight up and down (meaning straight push or pull movements). However, the data revealed that about half the measured movements were performed with different angles, which suggest that the participants did not move the mouse straight from one corner to the next. These random mouse movements in the learning phase, makes it difficult to know if the verbal cues and the physical arm movements overlapped as intended.

The mouse-tracking data revealed a problem in the research design making the motoric movements in the learning phase less consistent both within and between subjects. In general, it appears that the participants had too much freedom to move the mouse around during the course of the learning phase. If this experiment were to be replicated, a good idea would be to change the programming so that the mouse-cursor would have to remain in each corner of the screen until a new corner is highlighted. Thus, there will be no room for random mouse movements in between. In other words, to make sure that the participants are limited to more controlled mouse movements, as it is an important aspect of the experiment.

Conclusion

In the present work, I provided empirical support that may suggest a connection between verbally presented action-words and action initiation in response to relevant stimuli. This suggest that action-specific words are causally related to the facilitation of relevant actions. I was however unable to determine how this causal connection was formed, as my initial prediction failed to reach significance. A post-evaluation of the experiment itself reveals limitations with the design that can be improved upon in future research.

References

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences, 22*, 577-660.

Barsalou, L.W. (2005). Abstraction as dynamic interpretation in perceptual symbol systems. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories* (389-431). Carnegie Symposium Series, NJ: Erlbaum

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645. doi: 10.1146/annurev.psych.59.103006.093639

Barsalou, L.W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science, 2,* 716-724. doi: 10.1111/j.1756-8765.2010.01115.x

Barsalou, L.W., Simmons, W.K., Barbey, A.K., Wilson, C.D., (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*, 84-91. doi: 10.1016/S1364-6613(02)00029-3

Breuer, A.T., Masson, M.E.J., Cohen, A.-L., Lindsay, D.S. (2009). Long-term repetition priming of briefly identified objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 487-498. doi: 10.1037/a0014734

Fodor, J.A. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT press

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

Glenberg, A.M., Robertson, D.A. (1999). Indexical understanding of instructions. *Discourse Processes, 28*, 1-26. doi: 10.1080/01638539909545067

Glenberg, A.M., Robertson, D.A. (2000). Symbol grounding and meaning: A comparison between high-dimensional and embodied theories of meaning. *Journal of Memory and Language 43*, 379-401. doi: 10.1006/jmla.2000.2714

Hauk, O., Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping, 21*, 191-201. doi: 10.1002/hbm.10157

Logan, G.D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22*, 1–35. doi: 10.1016/0010-0285(90)90002-L

Martiny-Huenger, T., Martiny, S.E., Parks-Stamm, E.J., Pfeiffer, E., Gollwitzer, P.M. (2017). From conscious thought to automatic action: A simulation account of action planning. *Journal of Experimental Psychology: General, 146*, 1513-1525. doi: 10.1037/xge0000344

Peirce, J.W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*. doi: 10.3389/neuro.11.010.2008

Pfeuffer, C.U., Pfister, R., MoutsoPoulou, K., Waszak, F., Kiesel, A. (2017) The power of words: On item-specific stimulus-response associations formed in the absence of action. *Human Perception and Performance, 43*, 328-347. doi: 10.1037/xhp0000317

Pulvermüller, F., Shtyrov, Y., Ilmoniemi R.J. (2003). *Neuroimage, 20*, 1020-1025. doi: 10.1016/S1053-8119(03)00356-2

Pulvermüller, F., Hauk, O., Nikulin, V.V., Ilmoniemi, R.J. (2005). Functional links between language and motor systems. *European Journal of Neuroscience, 22*, 793-797. doi: 10.1111/j.1460-9568.2005.03900.x

Smith, E.E., Kosslyn, S.M. (2014). *Cognitive psychology: Mind and brain*. Harlow: Pearson Education Limited.

Stanfield, R.A., Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*, 153–156. doi: 10.1111/1467-9280.00326