Lending the Brain a Hand

The Influence of Motor Activation on Language Processing and Language Learning

Jacqueline A. de Nooijer

Lending the Brain a Hand

The influence of motor activation on language processing and language learning

Het brein een handje helpen

De invloed van motoractivatie op taalverwerking en taal leren

Proefschrift

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam

op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 10 september om 15.30 uur

door

Jacqueline Angelique de Nooijer

geboren te Vlissingen

Erasmus University Rotterdam

Promotiecommissie

Promotoren

Prof.dr. R.A. Zwaan

Prof.dr. F. Paas

Prof.dr. T.A.J.M. van Gog

Overige leden

Dr. K. Dijkstra

Prof.dr. J.W. van Strien

Prof.dr. H. Bekkering

The research presented in this dissertation was carried out at the Erasmus University Rotterdam and was funded by the Netherlands Organisation for Scientific Research and National Initiative Brain & Cognition (project number: 056-33-016)

Cover design by J.A. de Nooijer
Layout by J.A. de Nooijer
Printed by GVO drukkers & vormgevers B.V.

ISBN: 978-90-6464-887-8

Contents

▪ Chapter 1

General introduction

Parts of this chapter also appeared in: De Nooijer, J. A., & Willems, R. M. (in press). What can we learn about cognition from studying handedness? Insights from cognitive neuroscience. In N. Hagemann, B. Strauss, C. MacMahon, & F. Loffing (Eds.), *Laterality in Sports: Theories and Applications.* San Diego, California: Academic Press, Elsevier.

There is an old saying that goes: 'When in Rome, do as the Romans do'. So, when I lived in Italy several years ago, I tried to do exactly that. For this to happen, I also had to immerse myself in the Italian language. I went to an Italian course to learn the basics, I talked about the weather with my Italian roommates and I even tried to talk to government officials in Italian (who of course laughed at my attempts). But I soon realized that no matter how well I would learn the vocabulary and grammar, I might not master one crucial aspect of the way Italians talk: They speak as much with their hands as they do with their mouths.

Although there might be cultural differences in the number or type of gestures that people use (indeed, research has shown that Italian children gesture more than American children; Iverson, Capirci, Volterra, & Goldin-Meadow, 2008), we all use our hands when we talk. When we are mad we throw our hands in the air and when we show someone the way, we point in the right direction. Gestures are so much part of our language that we are even hindered in what we want to say when we are (for whatever reason) prevented from gesturing (e.g., Stevanoni & Salmon, 2005). And we even gesture when we are on the phone and the person we are talking to cannot possibly see our hands (e.g., Bavelas, Gerwing, Sutton, & Prevost, 2008). That we use our hands when nobody can see the meaning of these movements, suggests that gesturing not only seems important for our communication, but also seems to serve a cognitive function, helping us think and solve problems (for discussion of evidence that this is the case, see: Alibali, Spencer, Knox, & Kita, 2011; Chu & Kita, 2011; Kita, 2000; Pouw, De Nooijer, Van Gog, Zwaan, & Paas, 2014; Thomas & Lleras, 2009; Wesp, Hesse, Keutmann, & Wheaton, 2001).

Findings regarding gesturing as an aid to thinking and problem solving also suggest that we might put the body to good use for improving memory or learning. Indeed, several studies have found beneficial effects of gesturing on understanding math problems (e.g., Broaders, Cook, Mitchell, & Goldin-Meadow 2007; Cook, Yip, & Goldin-Meadow 2012; Goldin-Meadow, Cook, & Mitchell 2009; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001) or problem solving (Alibali et al., 2011; Chu & Kita, 2011; Garber & Goldin-Meadow 2002). To illustrate, when children are presented with an equation, like 6 + 3 + 4 = __ + 4, while pointing with the index finger and middle finger to the 6 + 3 and then pointing with their index finger to the blank space, indicating they must add 6 + 3 to solve the problem, they were better at solving the problems than when they did not gesture (Goldin-Meadow et al., 2009). When encouraged to use gestures during a mental rotation task, performance increased compared to when gestures were allowed (but not encouraged) or prohibited (Chu & Kita, 2011). In addition, gestures not only affect math learning or learning to solve problems, but also language learning, which is the focus of this dissertation.

Using gestures to improve word learning

Vocabulary learning is an important part of language learning, whether it is the first language or a second language that has to be acquired (e.g., Folse, 2004). Improving ways to teach children and adults novel words could therefore greatly benefit language learners. Several ways to improve word learning have been the topic of investigation. For example, using the context in the form of concept wheels (Vacca & Vacca, 2002), or semantic word maps (Schwartz & Raphael, 1985), learning in combination with pictures (Sadoski, 2005) and using imagery methods (e.g., Atkinson, 1975; Cohen & Johnson, 2011) have been investigated as possible methods to improve word learning. Another method to improve vocabulary learning is by using gestures during learning (for a review see Hald, De Nooijer, Van Gog, & Bekkering, 2015). Before discussing *how* gestures could play a role in word learning, it is important to know *why* the use of the body could enhance performance on learning novel words, and this is explained by theories of *embodied cognition*.

Embodied cognition

According to embodied cognition theories, cognitive processes like language, thought and perception cannot be separated from how our body functions (Barsalou, 1999; Glenberg, 1997). In other words: our body influences our cognition. The influence of the motor system on cognition has been found in many cognitive domains. For example in memory tasks, where autobiographic memories are recalled faster and are remembered better over a two week interval when the participant is currently in the same body position as when the memory was created (i.e., when lying back in a chair it is easier to remember a visit to the dentist's office;

Dijkstra, Kaschak, & Zwaan, 2007). Also when having to give estimations of, for instance, the height of the Eifel tower, lower estimations are given when participants are unknowingly leaning to the left (Eerland, Guadalupe, & Zwaan, 2011). The idea here is that people in Western countries represent numbers on a mental number line where lower numbers are represented on the left side of space, while larger numbers are positioned on the right side (i.e., in countries where people read and write from right to left, this number line seems to be reversed; Zebian, 2005). Thus, leaning to the left activates smaller numbers more than larger numbers, affecting the estimations. Lastly, in processing emotions, the motor system also seems to play a role. For instance, sentences describing a person who frowns are read more slowly when frowning muscles are paralyzed and the emotion can, therefore, not be simulated (Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010).

The influence of the motor system on cognition is also found in language processing. As mentioned before, according to theories of embodied cognition, the way we understand the world is grounded in action and perception, not in abstract symbols (e.g., Barsalou, 1999). Word meanings that are stored in the mental lexicon, therefore, consist not only of semantic information but also of motoric information. When encountering a word, it is therefore suggested that mental simulations are created that are "the reenactment of perceptual, motor, and introspective states acquired during interaction with the world, body, and mind'' (Barsalou, 2008, p.618), which results in pre-motor activity (e.g., Willems, Toni, Hagoort, & Casasanto, 2010). Thus, language and action are closely connected (e.g., Fischer & Zwaan, 2008; Willems & Casasanto, 2011; Willems & Hagoort, 2007). Evidence for this phenomenon comes from studies that show that hearing an action word, such as *to kick,* activates the brain's motor and premotor cortex, which are also activated when actually executing this action (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Tettamanti et al., 2005). The reverse also seems likely: if language can influence the motor system then activating the sensorimotor system can also influence language processing. Indeed, research has shown that action words that involve arm or leg movement (e.g., write, kick) were recognized faster after arm or leg activation (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005; see also: Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008).

From an embodied view on word learning, one would, therefore, expect that the use of gestures would activate the motor system (in the form of mental simulations), which might improve learning by reducing working memory load and/or enriching the mental representation of the word. As word meanings are also in part made up of motor information, using gestures that are congruent with what has to be learned and linking the novel words to the learner's congruent motoric experiences could lead to deeper understanding of word meaning (i.e., richer encoding and better consolidation of the new word). These ideas are also in line with the cognitive theory of multimedia learning (e.g., Mayer, 2014), which suggests that information from different modalities is processed via different working memory channels. Receiving both verbal and motoric information (in the form of gestures) that are processed by different channels, not only leads to the creation of a richer mental representation of what has to be learned, but also makes optimal use of different working memory subsystems (Baddeley, 1999). Working memory capacity is limited, and learning new information often imposes high working memory demands, but using different working memory modalities can reduce the load to some extent. Gestures might, therefore, reduce working memory load and/or create a richer mental representation of what has to be learned (e.g., Goldin-Meadow, 2010), although these explanations are not mutually exclusive. Combining verbal information with gestures might, thus, result in the creation of a richer mental representation that can facilitate the consolidation of the novel word. Activating the sensorimotor system through gestures during word learning can be accomplished in three different ways: observation, observation plus imitation, or enactment. The effects of using gestures during word learning will be discussed for each of these methods separately.

Observation of gestures

Observing a gesture that is congruent with what has to be learned (e.g., for the word 'book' a gesture could be: holding the hands next to each other, palm up, as if holding a book) can support word learning (e.g., Goodwyn, Acredolo, & Brown, 2000; McGregor, Rohlfing, Bean, & Marschner, 2008; Tellier, 2005). One explanation for this beneficial effect of gesture observation on learning comes from neuroscientific research that has shown that observing actions executed by others activate parts of the motor system that would also be activated when the learner would execute this action him- or herself. Some suggest that this is due to activation of the 'mirror neuron system' (Rizzolatti & Craighero, 2004), and although this system is highly debated (e.g., Glenberg, 2011; Gallese, Gernsbacher, Heyes, Hickock, & Iacoboni, 2011), such findings at least show that sensorimotor activation (which embodied accounts of cognition would predict to aid memory and learning, see prior section) can be accomplished via gesture observation.

The combination of observation of a meaningful gesture and verbal input seems to increase word learning already at a very young age: 24-months old children already gain larger vocabularies in their first language when their parents use both verbal information and iconic gestures during interaction (Goodwyn et al., 2000). In addition, in children of the same age, understanding of novel prepositions, like the word 'under' improves when congruent gestures are observed during learning (Mcgregor et al., 2008). Also in five-year olds, observation of gestures congruent with the meaning of the word seems to improve memorization of novel nouns and adjectives in the first language, compared to when no gestures were used (Tellier, 2005). In adults, observation of meaningful gestures can improve second language learning. For example, adults learned the meaning of novel Japanese verbs better when a gesture congruent with the meaning of the word was observed during learning than when no gestures were presented during learning (Kelly, McDevitt, & Esch, 2009).

Imitation of gestures

Gesture imitation has been shown to improve word learning in adults when concrete (Macedonia, Müller, & Friederici, 2011) and abstract (Macedonia & Knösche, 2011) words from an artificial language were learned while imitating iconic gestures. Moreover, when learning new expressions in a second language, learning performance improved when adults imitated congruent gestures, compared to when expressions were learned without seeing the gesture (Allen, 1995). Moreover, although the studies discussed above seem to suggest that word learning can be improved when the verbal definition is combined with observation of a gesture that is congruent with the meaning of a word, it can be expected that imitating an action leads to even higher performance scores than merely observing a gesture.

Both action observation and imitation activate the motor system, but action observation activates the motor system to a lesser degree than action execution (Keysers & Gazzola, 2009). When it is not sensorimotor activation per se, but the strength of the activation that matters, then imitating a gesture would be expected to lead to better learning. Indeed, it appears that word learning improves when children imitate a gesture compared to when they only observe a gesture, when learning words from a first (Tellier, 2005) or second language (Tellier, 2008). Brain research on recognition of words that were previously learned supports the idea that it is especially helpful for word learning when actions are actively performed (James & Swain, 2011; see also: Macedonia et al., 2011). fMRI research with children as participants suggests that motor areas of the brain were activated to a larger extent when hearing verbs that were previously learned by actively performing congruent actions, than when the actions were only observed (James & Swain, 2011). These findings indicate that sensorimotor information is reactivated when recognizing words that were previously learned, but only when having actively performed congruent actions. Similarly, when words are learned through imitation of iconic gestures during learning, brain activation in the premotor cortex was found when recognizing the new words (Macedonia et al., 2011). These studies also provide evidence for the idea that sensorimotor information may lead to a richer representation of what has to be learned, which could help to improve word learning (Wellsby & Pexman, 2014).

Given that very few studies have studied gesture imitation compared to gesture observation in word learning, its effectiveness should be further investigated. For example, most of these studies have applied imitation during encoding (immediately after observing the gesture that is congruent with the meaning of the verb; Tellier, 2005; 2008) when it can aid learning. Given the involvement of the motor system in recognition of learned words, gesture imitation might not only be helpful during encoding but also during retrieval (i.e., test taking). When a gesture is imitated during retrieval, it might provide learners with additional cues that could help retrieve the learned information.

Enactment

A third method of using gestures during word learning is via enactment. Enactment is defined here as self-generated action, without the presence of an example (in contrast to gesture imitation). This phenomenon of improved memory after self-generated actions is called the enactment effect (e.g., Dijkstra & Kaschak, 2006; Engelkamp, 1998; Nilsson, 2000). Self-involvement in the task might therefore lead to better word learning performance than merely imitating someone else's movement (e.g., Kormi-Nouri, 2000).

Indeed, there is some evidence that word learning in children can be improved via enactment. For example, when children were asked to show what a novel abstract verb, adjective or noun means through enactment, they learned more verbal definitions from their first language, than when no enactment was required (Casale, 1985). Even in the 60's, enactment was used in the classroom, as it was thought to enhance understanding and memory of novel words (Asher & Price, 1967). In adults, similar effects were found. In a longitudinal study adults learned words from an artificial language while only listening to the word, or by imitating a gesture performed by the experimenter. Both at a short-term interval and after 14 months, words that were enacted were remembered better (Macedonia, 2003). Lastly, memory for both concrete and abstract words from the first language is improved when adults are cued during recollection with gestures that they made for the to-be-remembered words, compared to when they were not cued with gestures, or with someone else's gestures (Frick-Horbury, 2002).

However, the effects of enactment might not always be stronger than the effects of action observation. To illustrate, reading action phrases (e.g., peel a potato) in combination with watching a congruent gesture or producing a mime does not lead to a difference in the number of action phrases that are correctly recalled (Feyereisen, 2009).

Factors determining the effectiveness of gestures during word learning

The studies described above show that gesture observation, imitation, and enactment might improve word learning because of the link that exists between action (i.e., the motor system) and language. However, the effects of gesture observation, imitation, and enactment on word learning have not yet been directly compared. Moreover, the effectiveness of these gesture conditions might be dependent on the type of word that has to be learned, as there are some indications that only learning of concrete verbs improves performance under gesture conditions (for a review see Hald et al., 2015).

Age is another factor that might potentially affect the effectiveness of gesturing, and especially for action verb learning. Although word learning seems to improve in adults when gestures are used during learning (e.g., Kelly et al., 2009; Macedonia & Knösche, 2011; Macedonia et al., 2011), there are reasons to believe that the effects of these gestures might be less strong than for children. One of these reasons is that the motor information conveyed through gestures is more relevant for children, as they do not have motoric experience with many actions yet, meaning that the definition of the word as well as the action it defines are new to them. Adults who are learning a second language, only have to link a novel word (i.e., label) to an already known action word from their first language, and adults learning definitions of unfamiliar action words in their first language are likely to have seen the actions being performed even though they may not have known the word for it. This would seem to make the gesture more redundant and the learning task easier for adults. Previous research has shown that gestures are mostly effective in more difficult tasks (McNeil, Alibali, & Evans, 2000) and when the gestures are relevant for learning (Kiefer & Trumpp, 2012). The effects of gesture observation and imitation have, however, never been directly compared to each other in children and adults, and so it is unclear whether there is a difference in effectiveness for word learning between these two groups when learning new words in combination with gesture observation or imitation.

Additionally, the effectiveness of using gestures during word learning might be influenced by the working memory capacity of the learner. If a richer mental representation is created when using gestures during learning, and that representation can lighten the working memory load, then learners with a larger working memory capacity might not benefit as much from using gestures during learning. For them it might not be necessary to lighten the working memory load. Given that working memory capacity plays a role in language acquisition (Ellis & Sinclair, 1996), this could mean that the use of gestures during word learning has a different effect on people with high and low language abilities.

Lastly, a factor that seems to play a role in whether gestures are helpful in learning words is the meaningfulness of the gestures. Most studies in this field suggest that if gestures are used in word learning they should be meaningful. For example, in a study where adults learned words from an artificial language in combination with observing and imitating meaningful iconic gestures or non-meaningful gestures, the use of meaningful gestures led to better memory performance (Macedonia et al., 2011). Also, when teaching adults the meaning of Japanese nouns and verbs by means of congruent and incongruent gestures, participants learned more words combined with congruent gestures, suggesting that gestures have to be meaningful in order to have an effect on word learning performance (Kelly et al., 2009). This seems logical as only a richer representation of what has to be learned can be created when the verbal and motoric information (in the form of gestures) convey the same information. Another explanation for why only meaningful gestures can improve word learning could be that non-meaningful gestures interfere with the learning process. Mental simulations are created when hearing a word or a definition, and seeing an action or object (or a picture thereof) can also automatically activate motor information (Borghi et al., 2007; Sumner et al., 2007). Therefore, when seeing a non-meaningful gesture during learning, the motor information conveyed through the gesture does not match the motor activation evoked by the definition of the word. This mismatch might, therefore, hinder the encoding and consolidation process. Similarly to using incongruent gestures during word learning, seeing actions performed by someone who has a different hand preference (i.e., seeing a left-hander perform an action as a right-hander), could also interfere with learning new words. To understand why this might be so, first a more thorough explanation of mental simulations is given.

Mental simulations of actions during language processing and learning

As mentioned above, the motor system is involved in processing language, which means that when we understand language, we activate our motor system. According to embodied cognition theories this (pre-)motor activation is reflected in mental representations or mental simulations that are created when we hear or read something. To illustrate, when hearing a sentence such as "Close the drawer," (which implies a movement away from the body) the denoted action interacts with the subsequent performance of an action, meaning that, in this example, a movement away from the body is facilitated (e.g., Glenberg & Kaschak, 2002).

Such findings suggest that we automatically make mental simulations of what we read and hear in a large array of situations, and that these mental simulations might even be necessary for many language processes. However, a new view of cognition, the pluralist view, suggests that the creation of mental simulations is not always required. According to this view, cognition involves both grounded and abstract symbols (e.g., Dale, Dietrich, & Chemero, 2009; Dove 2009; Zwaan, 2014). With the emergence of this view a discussion has started on when we need, or use, abstract or grounded symbols (Zwaan, 2014). Several factors could influence the extent to which mental representations are used during language comprehension. Whether we make a mental representation depends, for example, on the processing goal (Zwaan, 2014), but also on the experience we have with a certain action. However, research does suggest that at least in certain situations mental simulations are created (Glenberg & Kaschak, 2002; Zwaan et al., 2002). An interesting way to investigate whether premotor activation in the form of mental simulations can influence learning is by combining verbal information with the observation of pictures in which an action is executed by a right- or lefthander. Left- and right-handers structurally perform certain actions differently. According to the body-specificity hypothesis discussed below they might, therefore, also create different mental simulations.

The body-specificity hypothesis

According to the body-specificity hypothesis reading or hearing an action leads to the creation of a body-specific mental simulation of this action (Casasanto, 2009). For instance, when passively viewing the letters of the alphabet, right-handers activate part of the left premotor cortex (that is also active when writing with the right hand; Longcamp, Anton, Roth & Velay, 2003). When a similar study was conducted with left-handers, the right premotor area was activated (Longcamp, Anton, Roth & Velay, 2005), suggesting that when merely viewing letters, mental simulations are made of writing those letters. In addition, when left- and right-handers are asked to imagine (simulate) executing a number of action verbs, such as 'writing', right-handers activate left-premotor areas (congruent with activation of the right arm) while left-handers activate right premotor areas (congruent with activation of the left arm; Willems et al., 2010). There might even be a causal link between the premotor cortex and understanding of action verbs. This hypothesis seems to be supported by the finding that right-handers are faster to indicate that a given manual action (e.g., to write, to throw) or non-manual action word (e.g., to earn, to wander) is an existing verb after stimulation of the left compared to the right premotor cortex (Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). People seem to understand actions relative to how they would perform the action with their own body.

Left- and right-handers show a different pattern of motor activity not only when imagining executing actions but also when merely listening to sounds of tools, this seems to be the case. When presented with sounds of tools that can be manipulated with one hand (i.e., hammer, saw, screwdriver) versus sounds that animals make, fMRI data showed that when hearing tool sounds right-handers activate left-hemisphere motor-related areas, while left-handers activate right-hemisphere motor areas (Lewis, Phinney, Brefczynski-Lewis, & DeYoe, 2006). This difference did not exist for the sounds of animals. These results suggest that tool sounds are learned in the context of how they are used (with either the left or right hand). In this case, motor imagery might have facilitated the recognition of the tool sounds.

In other words, information might be processed differently by left- and right-handers, in line with differences in how they use their body, but would this in any way affect memory and learning? Does the way we interact with an object influence how well we can remember it?

The influence of mental simulations on memory

According to the body-specificity hypothesis word meanings are in part made up of the simulations of one's own actions (Casasanto, 2009). Mental simulations might therefore be dependent on the experiences people have with an action or object. These experiences might also influence how well the objects or actions are remembered. For instance, words that we can manipulate (e.g., “camera”) are remembered better than words we do not interact with (e.g., “table”; Madan & Singhal, 2012). If whether we interact with objects can influence how well we remember the object, then the way we interact with it (i.e., their affordances, e.g., Gibson, 1977) might also influence our memory. This, indeed, seems to be the case. A study that investigated this issue showed left- and right-handed participants a 3 x 3 grid with cups around the grid. The handles of these cups were pointed either towards the right or towards the left. Participants were then given instructions as to how to move the cups across the grid. Half of the participants executed the memorized actions with their dominant hand, the other half with their non-dominant hand. Results showed that right-handers remembered more instructions when an object’s handle was oriented to the right and actions had to be performed with the right hand. This suggests that motor simulation can support sequential memory for action and that memory is embodied, or at least that if the handle is oriented in congruence with the use by a right-hander, the mental simulation might be facilitated, enhancing memory performance (Apel, Cangelosi, Ellis, Gosling, & Fischer, 2012, but see: Pecher, 2013; Pecher et al., 2013). It is unclear whether memorization of known action words would also be affected by seeing a picture that either matches or mismatches the participants’ handedness (i.e., seeing a picture of a right-hander or left-hander performing the action).

The influence of mental simulations on learning

As the focus of this dissertation lies on word learning, an important question is whether mental simulations could not only affect memory, but also influence word learning. Seeing a picture that mismatches the mental simulation created by verbal information might hinder the consolidation of new information, as motor information is part of the word’s meaning. Thus, receiving contradictory motor information might hinder learning new information. So, if we hear a definition of a word and then see a picture that either matches (i.e., seeing a picture of a right-handed perspective as a right-hander) or mismatches (i.e., seeing a picture of a left-handed perspective as a right-hander) the mental simulation created, this might hinder word learning.

Possible effects of mismatching mental simulations on word learning may be bound by the observer's perspective. Research suggests that the first-person perspective is more embodied than the third-person perspective (e.g., Lorey et al., 2009), as evidenced, for instance, by findings that imagining hand movements from a first-person perspective leads to stronger activation in motor areas of the brain than when imagining the same from a third-person perspective (Lorey et al., 2009; Ruby & Decety, 2001). Motor information, therefore, seems to play a bigger role in processing pictures from a first- than from a third-person perspective. If motor information is part of the meaning of a word as stored in the mental lexicon, which is suggested by the body-specificity hypothesis (Casasanto, 2009), then receiving contrasting information (seeing the left-handed perspective as a right hander) might hinder consolidation of the meaning of the word. However, if the third-person perspective does not activate motor information to the same degree, the consolidation process might not be hindered.

Influencing right- but not left-handers

If handedness would indeed influence the way we memorize and learn, the question is whether this effect would apply to both right-handers and left-handers. Studies have found an influence on performance when right-handers were presented with the left-handed perspective (in the form of a picture, of a handle oriented to the left), but not when left-handers were presented with the right-handed perspective. Left-handers are not affected (or less affected) by seeing the right-handed perspective (e.g., Apel et al., 2012). Also in other studies similar results are obtained. For instance, when asked to press a button in a reaction time experiment with either the right or left index finger or both fingers, right-handers showed different brain activity (i.e., an increase of activity in dorsal premotor and right primary sensorimotor cortices) for making a unilateral versus a bilateral response, while left-handers did not. This suggests that there are differences in bimanual motor control related to handedness (Klöppel et al., 2007). Another study showed that perceived distance of objects varies when an object is in a difficult versus an easy to grasp position, but only in right-handers. Right-handers perceived tools as farther away when they are difficult to grasp, but left-handers did not (Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009).

Why is this? Why do right-handers seem to behave differently when presented with a left-handed perspective (or left-handed action), but the opposite is not true for left-handers? One explanation might lie in the fact that right-handers outnumber left-handers by far (i.e., only approximately 10% of the population is left-handed; Coren & Porac, 1977). Left-handers, therefore, live in a right-handed world. They use their right hand more often than right-handers use their left hand and might therefore not have very strong hand-specific embodied connections (e.g., Stins, Kadar, & Costall, 2001). For this reason, left-handers might not have strong associations between objects and actions, as they need to be more flexible in the use of many objects. Right-handers are, however, influenced by observing a handedness perspective that (mis)matches their own. An interesting question is whether the mismatch in perspective that right-handers seem to experience, cannot only be found in behavioral tasks, measured by reaction times, but whether early differences in brain activity can be found to indicate a difference in the processing of the left- and right-handed perspective. This could provide insights as to where these effects stem from.

Overview of the studies presented in this dissertation

The present dissertation is concerned with the following research questions: What gesture instructions (observation, imitation, or enactment) are most effective for improving verb learning in primary school children, and are these gestures equally effective for learning verbs from different verb categories and for children with different levels of language ability? In addition, it is investigated whether the effectiveness of gestures in word learning is similar for children and adults. Lastly, this dissertation focuses on whether learning and memorizing action words are influenced by seeing depictions of actions that mismatch the mental simulation that is automatically created when hearing the definition of an action word. In addition, several studies are devoted to the question how these effects can be explained.

The studies presented in Chapters 2 and 3 investigate the effects of different kinds of gesture use on primary school children's word learning. The study reported in **Chapter 2** makes a systematic comparison of learning novel action verbs in combination with gesture observation (video), observation plus imitation, or enactment, compared to a no gesture control condition. The study described in **Chapter 3** focuses on the effects of the timing of gesture imitation (i.e., immediately after observing gestures on video or from memory prior to the test or both). Moreover, in both chapters, action verbs from different categories are used (i.e., object-manipulation, locomotion, or abstract). This allowed for investigating whether gesture type (observation, imitation, enactment) or timing (imitation immediately or delayed) have differential effects when learning verbs that have a strong connection to the motor system (e.g., manual action verbs) or verbs that have a minimal connection to the motor system (e.g., abstract verbs).

Chapter 4 presents a study investigating whether the effectiveness of gesture observation plus imitation compared to observation only when learning object-manipulation verbs differs between children and adults and differs as a function of whether the action is familiar or not. A factor that might play a role in whether gestures are indeed effective in word learning is whether the word of which the definition has to be learned is already known to the learner. When the word is known, only the precise definition of the word has to be learned, while if the word is still unknown, both the word and the definition have to be learned. In this case, the gestures provide additional information (and might therefore support learning), because if the word is completely novel, the learner does not have motoric experiences with the action yet (i.e., no automatic mental simulation of the defined action). As many studies on the effects of gestures on learning have been conducted with children, it is unclear whether the same strategies that are used with children can be used in adults. This study, therefore, focuses on the potential influence of age and of the novelty of a word in the effectiveness of imitation during word learning.

The experiments presented in **Chapter 5** take a different approach to studying the effect of motor activation on word learning, using static pictures of actions and adult participants to investigate how pictures that (mis)match participants' motor simulation affect learning of concrete action verbs. Participants had to learn the verbal definition of an action word in an artificial language (e.g., ‘luko’ means ‘to pour from a container’). First, they heard the verbal definition (which presumably activated a motor simulation of the action) and then, they heard the verbal definition again, while seeing a picture that either matched their mental simulation (i.e., right-handed pictures for right-handers) or mismatched their mental simulation (i.e., right-handers seeing a left-handed first-person picture perspective). In this chapter, it is also investigated whether the influence of (mis)matching mental simulations is different for right- and left-handers. As left-handers live in a right-handed world, they are used to performing some actions with their right hand as well as their left. This could result in them having less strong hand-specific embodied connections. They could, therefore, be less influenced by learning words in combination with seeing an action that mismatches their handedness perspective.

In the experiments in **Chapter 6**, the ideas of Chapter 5 are further investigated, but now the focus lies on the role of observation perspective. Is there only a role for mental simulations in word learning in pictures from a first-person perspective or is the same effect present when viewing third-person perspective pictures that mismatch the learner’s mental simulation? This is an interesting question because it will provide insight into the circumstances under which mental simulations can influence cognitive processes, such as word learning. Therefore in Chapter 6 it is investigated whether the effect that is expected in Chapter 5 for right-handers can be replicated and whether the same effect would be found with third-person perspective pictures and how learning words in combination with right-or left-handed picture perspectives relates to performance on words learned without any pictures. This could provide an answer to the question of whether mismatching mental simulations hinder learning or whether matching mental simulations support learning. In this chapter the procedure is the same as in Chapter 5, however in these studies only right-handers are used.

The experiments presented in **Chapter 7** take the results of Chapter 5 and 6 one step further. The first experiment investigates whether pictures that (mis)match right-handed participants’ mental simulation would also affect word memory (i.e., instead of learning, participants merely memorized a list of action verbs combined with either left- or right-handed pictures). The second experiment investigates whether participants' visual attention allocation could provide insight into the mechanisms underlying the mental simulation mismatch effect, by using eye tracking.

In the study presented in **Chapter 8**, a different method of investigation is used, namely EEG, to further investigate the nature of mental simulations based on (mis)matching verbal and pictorial information. In this study left- and right-handed perspective pictures were shown to the (right-handed) participants while they listened to a sentence that either matched or mismatched the content of the picture. The central question here is whether a mismatch in handedness perspective is also reflected in an ERP component, the N400. This component is an index of semantic processing that reflects the neural mechanisms of semantic integration into a context. This study could therefore provide insight into whether mental simulations play a role in the early process of the integration of information from the verbal and visual modality. Lastly, **Chapter 9** provides a summary of the main findings, along with a discussion of theoretical and practical implications and suggestions for further research.

▪ Chapter 2

Words in action: Using gestures to improve verb learning in primary school children

This chapter has been published as:

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2014). Words in action: Using gestures to improve verb learning in primary school children. Gesture*, 14,* 47–70. doi 10.1075/gest.14. 1.03noo.

Abstract

Research on embodied cognition has shown that action and language are closely intertwined. The present study seeks to exploit this relationship, by systematically investigating whether motor activation would improve eight-to-nine year old children's learning of vocabulary in their first language. In a within-subjects paradigm, 49 children learned novel object manipulation, locomotion and abstract verbs via a verbal definition alone and in combination with gesture observation, imitation, or generation (i.e., enactment). Results showed that learning of locomotion verbs significantly improved through gesture observation compared to verbal definitions only. For learning object-manipulation verbs, children with good language skills seemed to benefit from imitation and enactment, while this appeared to hinder children with poor language skills. Learning of abstract verbs was not differentially affected by instructional condition. This study suggests that the effectiveness of observing and generating gestures for vocabulary learning may differ depending on verb type and language proficiency.

Introduction

According to theories of embodied cognition, action and language are closely intertwined (e.g., Fischer & Zwaan, 2008). Not only does language processing affect the motor system, but the motor system also affects language processing. For instance, hearing action words such as 'kick' activate areas of the brain's motor and premotor cortex that are also active when executing the denoted action (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Tettamanti et al., 2005). In addition, research has shown that arm and leg activation results in faster recognition of action words involving arm or leg movements (Pulvermüller, Hauk, Nikulin, & Illmoniemi, 2005). In a similar vein, activation of the motor system by intentional action can facilitate the processing of functionally manipulable words (Rueschemeyer, Lindemann, van Rooij, van Dam, & Bekkering, 2010).

This tight link between action and language might be exploited to improve learning, by combining verbal instructions with motor stimulation. According to the embodied cognition framework, language is grounded in action (Barsalou, 2008). Considering that children are still in the process of acquiring their first language, they might have an added benefit of motor activation during learning. A meta-analysis of gesture studies supports this notion, showing that children are more likely to benefit from gestures during speech in understanding a message than adults and teens (Hostetter, 2011). In the current study gestures are defined according to the restricted definition of the Oxford English Dictionary as "movements of the body or parts of the body that are expressive of thought or feeling". Research has shown that gestures that convey part of the meaning of what has to be learned can foster learning. Not only making gestures (Ehrlich, Levine, & Goldin-Meadow, 2006), but also imitating (e.g., Cook, Mitchell, & Goldin-Meadow, 2008) or observing gestures (e.g., Ping & Goldin-Meadow, 2008) has been shown to foster acquisition of a variety of tasks.

Learning by observing, imitating, and making gestures

There are several ways in which verbs can be taught to children using gestures, one of which is to combine verbal input with action observation, that is, observing a demonstration by a human model of a gesture portraying the meaning of the verb, by means of dynamic visualizations such as video. Observational learning is a powerful form of learning, for a wide variety of tasks, especially when learners lack prior knowledge (for reviews, see Atkinson, Derry, Renkl, & Wortham, 2000; Bandura, 1986; Renkl, 2011; Van Gog & Rummel, 2010). Research in neuroscience has shown that the mere observation of actions performed by others activates those parts of the motor system that would also be involved in executing the action oneself (it is claimed that this involves the mirror-neuron system; Rizzolatti & Craighero, 2004). It has been suggested that this is the reason why dynamic visualizations that show human movement can improve learning of procedural motor tasks, whereas dynamic visualizations for other types of tasks are not very effective (for reviews, see Höffler & Leutner, 2007; Van Gog, Paas, Marcus, Ayres, & Sweller, 2009). Given the close relationship between the motor system and language, we hypothesize that gesture observation in dynamic visualizations might also facilitate learning of non-human movement tasks, such as verb learning. Some evidence for this assumption comes from a study showing that when receiving verbal instructions in combination with action observation, five year old children can memorize a list of concrete nouns and descriptive adjectives better when the words are accompanied by a gesture (Tellier, 2005). Action observation seems to benefit not only memorizing items but also applying new information (Ping & Goldin-Meadow, 2008).

What happens when we combine observation with imitation? Although there are indications that mere action observation facilitates learning, action execution has been found to enhance memory traces (e.g., Dijkstra & Kaschak, 2006; Engelkamp & Cohen, 1991). Learning might therefore improve even more when the learner’s own motor system is engaged, via action imitation. For example, sentences uttered in an unknown accent are better understood when listeners imitate the accent, than when they only hear, repeat, or transcribe the sentences (Adank, Hagoort, & Bekkering, 2010). Furthermore, imitation of observed gestures has an effect on memorization. Children can remember more items from a list with familiar concrete nouns and descriptive adjectives when imitating gestures that accompany the words than when only observing these gestures, an effect that was found both for words from the first language (Tellier, 2007) and for words from a second language (Kelly, McDevitt, & Esch, 2009; Tellier, 2008). Also, using an artificial language, action imitation was shown to lead to better recall of abstract sentences (Macedonia & Knösche, 2011). These findings suggest that gesture imitation might have a beneficial effect on learning the meaning of verbs over and above mere observation.

So, action observation and imitation might both activate the motor system, which, according to embodied theories, leads to improved learning, and this activation could be expected to be stronger for imitation than for observation. However, memory for actions also seems to improve when actions during encoding are self-generated, a phenomenon that is called the enactment effect (see review by Engelkamp, 1998; Nilsson, 2000). Studies investigating memory for action events have typically compared recall or recognition for action phrases that were only read and action phrases that were enacted, in which memory was found to be superior for the latter (e.g., Dijkstra & Kaschak, 2006; Engelkamp, Zimmer, Mohr, & Sellen, 1994). For example, when children were asked to show what a novel abstract verb, adjective, or noun means through action execution, after having heard the verbal definition, they learned more words than when this action execution component was not required (Casale, 1985). However, the effect of self-generated action might not be larger than the effect of action observation. When participants were asked to read phrases containing action verbs or were additionally required to watch a gesture or produce a mime, a facilitative effect on memory was found for both action observation (i.e., watching a gesture) and enactment (i.e., producing a gesture). There was, however, no significant difference between these conditions (Feyereisen, 2009).

The present study aims to investigate the effectiveness of these different types of instruction (i.e., action observation, action imitation, enactment) for verb learning. Verbs are especially interesting for studies on gesturing, given the existence of verbs that have a direct link to the motor system (e.g., to stride, to chisel) and verbs that do not have a direct link to the motor system (e.g., to recruit).

Gestures and verb learning

Most studies that have investigated the role of the motor system in understanding language using verbs, have used object manipulation verbs or functionally manipulable verbs, (i.e., manual verbs such as, to throw, to write; e.g., Ruschemeyer et al., 2010; Tessari, Canessa, Ukmar, & Rumiati, 2007; Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). Object-manipulation verbs are, however, just one class of action verbs, and it is an open question whether gesturing would be beneficial for learning other types of action verbs. Locomotion verbs (e.g., to stride), for example, form a category of action verbs, that has –to the best of our knowledge- not been investigated yet. These verbs can be distinguished from object-manipulation verbs in a number of ways.

First of all, whereas object-manipulation verbs are mostly manual verbs, locomotion verbs require leg activation. These verbs therefore activate different parts of the motor cortex (e.g., Hauk et al., 2004; Tettamanti et al., 2005; see also: Hickok, 2010). Furthermore, although both object-manipulation and locomotion verbs can be goal-directed, locomotion verbs do not always denote intentional goal-directed activities (e.g., to stroll). Research has shown that the mirror neuron system (MNS; Rizzolatti & Craighero, 2004), is more active for goal-directed actions than for non-goal-directed actions (Nyström, Ljunghammar, Rosander, & von Hofsten, 2011). In sum, these findings suggest that motor activation due to observing or making gestures during verb learning could be more beneficial for object-manipulation verbs than for locomotion verbs, because object-manipulation verbs show stronger MNS activation because of their goal-relatedness.

Both object-manipulation and locomotion verbs are action verbs and have an obvious connection to the motor system, however, abstract verbs have not. Whether abstract verbs will also activate motor areas is not clear (Borghi, 2005). Although action verbs elicit more activation than abstract non-motor verbs in motor areas of the brain (Van Dam, Rueschemeyer, & Bekkering, 2010), the motor system influences both concrete and abstract language (Glenberg et al., 2008). According to the embodied cognition framework, abstract concepts are also grounded in interactions we have with the world. For instance, given that the abstract concept of time can be understood in terms of spatial experience (Boroditsky & Ramscar, 2002), abstract concepts could refer indirectly to bodily experiences (Borghi, Caramelli, & Setti, 2004). If this assumption is correct, abstract words should benefit from gesturing during verb learning as well.

In sum, there are different ways in which gestures can be used in instruction, and given that there are few direct comparisons, and –to the best of our knowledge- none involving all three types of gesturing (observing, imitating, and generating), the question, is which of these methods is the most effective. Moreover, studies on gesturing and language learning have mostly addressed second language learning, and have mostly used functionally manipulable words that have a concrete link to the motor system (e.g., Ruschemeyer et al., 2010; Tessari et al., 2007;) very little research has considered abstract verb learning and first language learning. Therefore, another question is whether the effects on learning of those different ways of incorporating gestures in instruction, are the same for different kinds of verbs (concrete, abstract) in the first language. In addition, studies investigating the role of gestures in word learning typically use free recall tasks (e.g., Tellier, 2007, 2008) or recognition tasks (e.g., Kelly, Devitt, & Esch, 2009). From an educational perspective, it is interesting to know whether gestures cannot only help children memorize new words better, but also whether they help them understand the meaning of those words better.

What we should also consider is that learning words by using observation, imitation, or generation of gestures, might not be equally beneficial for all learners. For example, a study with adults showed that gesturing aided recall of both concrete and abstract words more for adults with low verbal skills than for adults with high verbal skills (Frick-Horbury, 2002). Verbal skills differences were also reduced when imagery was used during recall (Ernest, 1977). When a gesture is encoded during word learning, it might later act as an additional cue for memory, and given that differences between people with high and low verbal skills are associated with differences in cognitive processing (e.g., retrieval), an additional cue for retrieval of information in the form of gestures might be especially beneficial for people with low verbal skills. Individual differences in verbal ability might also differentially affect children's learning in the different gesturing conditions.

Hypotheses

To try to answer the research questions described above, we conducted an explorative study that aims to systematically investigate to what extent motor activation, in the form of gesturing, can improve eight-to-nine year olds' learning of different types of verbs in the first language, using both a recall and comprehension test.

Based on the above review of the literature, we can formulate several hypotheses concerning the effect of the different instructional conditions (i.e., verbal definition, action observation, action imitation, enactment) and the effect of verb type (i.e., locomotion, object manipulation and abstract). Regarding the differences between the conditions we hypothesize that the number of verbs learned in the verbal definition condition would be lower than in the three other conditions in which gestures are used (hypothesis 1). Moreover, it can be hypothesized that action execution (i.e., both action imitation and enactment) would be more effective for learning than action observation (hypothesis 2) based on studies that found that action execution enhances memory traces and facilitates recall (e.g., Cohen & Otterbein, 1992; Engelkamp & Cohen, 1991). Finally, although self-generated action (i.e., enactment) has been found to improve memory for known items (e.g., Engelkamp et al., 1994), not all studies have found this effect (e.g., Feyereisen, 2009), and it can be questioned whether it would improve learning new verbs more than imitation, as generating a gesture for a known word that needs to be remembered is very different from generating a gesture for a word that you are learning for the first time. Therefore, we cannot formulate a clear hypothesis for the difference between imitation and enactment (question 1).

Concerning the different types of verbs, the effect of motor activation on learning of abstract words can be expected to be significantly smaller than the effect on locomotion and object manipulation verbs (hypothesis 3). Action words and words denoting manipulable objects reliably activate the motor system more than abstract words (Van Dam et al., 2010) and therefore the effect of motor activation on learning abstract verbs might be less pronounced. The difference in prediction for the locomotion and object manipulation verbs is less clear-cut. As mentioned before, the object manipulation verbs are more goal-directed than the locomotion verbs and the mirror neuron system appears to be more active during the observation of goal-directed actions than during the observation of non-goal directed actions (Nyström et al., 2011). Therefore, the gesturing condition might result in stronger motor activation and thus better learning for object-manipulation verbs than locomotion verbs (hypothesis 4).

Further, a difference in effectiveness of the use of gestures during learning for children with high and low verbal skills is expected, in that children with low verbal skills can be expected to benefit more from gestures than children with high verbal skills, because of the additional memory cue that gestures provide them (hypothesis 5). Lastly, there might be interaction effects between method of instruction and verb type. It could for instance be expected that abstract words are learned less well in the enactment condition than locomotion or object manipulation words, because it is more difficult to generate a gesture that is congruent with the meaning of an abstract word. However, because no other study has systematically investigated the effect of different types of gesture instruction on learning different verb types, we cannot formulate clear hypotheses on this matter on the basis of the literature (question 2).

To summarize, for the different gesture conditions we expect the following performance pattern: verbal definition < gesture observation < gesture imitation & enactment. For the verb types we expect the following pattern: abstract < locomotion < object-manipulation. Finally, we expect the use of gestures to be more beneficial for children with low verbal skills than for children with high verbal skills.

Method

Participants

Fifty-three Dutch primary school children participated in this study. The children (18 male, 31 female) were between eight and nine years of age (*M*= 8.6, *SD*=0.6) and were recruited from two primary schools in the same area of the Netherlands. Four children were excluded from data analysis because they were bilingual, so the analyses are based on forty-nine participants.

Materials

Verbs & Gestures. The stimuli consisted of 24 verbs in total, distributed over three sets of eight verbs, with each set representing a different verb type: locomotion (e.g., to stride), object-manipulation (e.g., to chisel) or abstract (e.g., to recruit). The verbs from these different categories were matched for phoneme length and for lexical frequency by determining the mean logarithmic lemma frequency per million for each verb type using the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). Furthermore, the level of concreteness was assessed for each category via the MRC psycholinguistic database (Coltheart, 1981). This database only contains English words, so the English translations of the Dutch verbs were used to assess the level of concreteness. Because not all verbs appeared in this database, the use of this measure was a mere indication of the concreteness of the verbs in each category. The mean concreteness scores per verb type suggest that the object-manipulation verbs are the most concrete, followed by the locomotion verbs and, as expected, the abstract verbs were the least concrete. To minimize the chances that children were familiar with the verbs that they were taught during the experiment, the Measure of Lexical Richness (MLR) was used. The MLR is a database derived from words that children are most likely to come into contact with at primary school. These words are divided over nine word lists (Schrooten & Vermeer, 1994), which are ordered in terms of lexical frequency. List one contains the highest frequency words, which children acquire relatively early, while list nine contains the lowest frequency words, which are learned in the upper grades of primary school. The verbs in this experiment only appeared on word lists eight or nine, or they did not appear on these lists at all, meaning that these words are most likely not learned before age twelve. This method minimized the chance that the children knew the verbs prior to the experiment.

For each verb a video of a congruent gesture was created, performed by a young adult actor. All actions were depicted as realistically as possible and could be described as pantomimes. The object-manipulation verbs were manual, the locomotion verbs involved leg movement (e.g., seeing someone striding), and for the abstract verbs either hands, legs, or both could be used. For example, for the abstract verb ‘to dismiss’ (defined as ‘to send someone away’), the children saw a person making a “shoo” gesture, waving the hand in a forward manner. For a complete list of the verbs and the gestures see the Appendix.

Pilot testing. All material was tested in two pilots. The first pilot test aimed to test the complexity of the verbs and the degree of difficulty of the posttest. The definitions of *twelve* verbs per verb type (i.e., locomotion, object manipulation, abstract) were read twice to 32 Dutch eight-to-nine year olds in blocks of six, and they were instructed that they should try to learn these definitions. After six verbs, the children were asked to provide the definitions of the newly learned words and to fill out a fill-in-the-gap test, which contained six sentences. Each of the six novel verbs had to be used once. The reason for using this pilot test was to examine if all the sentences it contained were understood by the children and to investigate if the verbs were either too difficult or too simple. Because the verbal definition condition is predicted to have the smallest positive effect on learning compared to the gesture conditions, we used these performance data to select verbs that showed evidence of learning while still leaving room for improvement on the test across conditions (see the final paragraph of this section).

The second pilot study was conducted to ascertain that the gestures that were shown in the videos, gave an accurate representation of the meaning of the verbs. Twenty-two Dutch adults were asked to match 36 verbs to gestures shown in videos. The experiment consisted of six blocks. In each block a list of six verbs was presented, after which the participants saw six gestures that they had to match to one of the six verbs. Furthermore, participants were asked to score the gestures on their compatibility with the verbs on a 7-point scale, where 1 meant that the gesture was a poor fit for the verb and 7 meant the gesture was an accurate portrayal of the verb's meaning.

On the basis of the combined results of the two pilot tests, *eight* verbs were chosen per verb category to be used in the experiment. The verbs on which children scored either extremely low or extremely high were removed, as well as those verbs that were not matched correctly to the gesture by at least 70% of the adults or that were not rated 5 or higher on the 7-point scale. There was no significant difference in how the quality of gestures from the three verb categories was judged on this scale.

Instructional conditions. The children learned the verbs under four different conditions: the verbal definition, action observation, action imitation, and enactment condition. In the *verbal definition condition* children were offered a verbal definition for a novel verb twice. The child only had to listen to the definition and try to remember it. In the *action observation condition* children heard a verbal definition twice and during the second presentation they were shown a video in which a gesture congruent with the meaning of the novel verb could be seen. The *action imitation condition*, was similar to the observation condition, but children were instructed to imitate the gestures they observed. Finally, in the *enactment condition* the children were instructed to make up a gesture during the second presentation of the verbal definition.

The children received a practice trial (with an unrelated word) before the start of each condition. In each condition, the children were told that they would be learning new words and that they should try to remember the new words and their meaning. A schematic presentation of the components of each condition is shown in Figure 1. Verbal definitions and gestures were video-taped beforehand to keep the presentation of this input equal for all children. An adult female gave the verbal definitions, while an adult male was shown to execute the gestures (with the definition spoken by the female as voice over) for the observation and imitation conditions.

Posttest. The post-test consisted of two parts. The first part was a recall test, in which children were asked to provide a definition of the six verbs of the test phase. The task was a paper-and-pencil test, but the target words were read orally to the children, to prevent any difficulties with encoding a written word from interfering with the learning capability of the child.

Given that a recall test does not measure comprehension of the verb, the second part of the post-test consisted of a fill-in-the-gap test. This test contained one sentence for each of the previously learned verbs, which resulted in six sentences per condition. No list of possible answers was given to prevent the test from turning into a recognition task. An example of a sentence from the fill-in-the-gap test was: “The carpenter wanted to build a closet and took a piece of wood and started to…. (chisel).” All correct answers were infinitives, given that conjugations would add a degree of difficulty.

Procedure

Before the experiment, scores on a standardized vocabulary test were collected from the schools for all children participating in the study. This is a validated test that is administered yearly by every primary school in the Netherlands (Cito, 2009). Scores on this vocabulary test consist of five categories, which have cut-off scores based on a national norm group: A (i.e., score in the top 25% of scores nationally), B (i.e., score in the next 25% range, just above the national average), C (i.e., score in the next 25%, just below the national average), D (i.e., score in the next 15%, scoring well below the national average), E (i.e., score in the 10% lowest scores).

The experiment was run in individual sessions on two consecutive days. On day one, children learned verbs under the verbal definition and action observation condition; on day two, they learned verbs under the action imitation and enactment condition. Each condition contained two verbs of each of the three verb types (i.e., locomotion, object manipulation, abstract) which resulted in six verbs per condition.

Verbal definition　　Action observation

Figure 1. Still frames of the female model reciting a verbal definition and the male model executing the gesture for ‘to chisel’.

The order of condition was fixed to avoid any carry-over or learning effects. If the gesture conditions had been presented before the verbal condition, children might have used mental simulation in the verbal condition, which is “the reenactment of perceptual, motor, and introspective states acquired during interaction with the world, body, and mind’’ (Barsalou, 2008). This would minimize the difference between the conditions and create a confound. However, the verbs were counterbalanced across conditions so that any learning effects could not be a consequence of the same verbs always occurring in the same condition. The posttest for the verbs in a particular condition was administered immediately after the learning phase.

Scoring & analysis

Both the recall test and the fill-in-the-gap test were scored by awarding one point for each correct answer. In the recall test children could also be rewarded a half point when one element of the definition was missing (e.g., ‘to hit a sharp object’ instead of ‘to hit a sharp object with a hammer’ for ‘to chisel’). A 4 X 3 repeated-measures analysis of variance (ANCOVA) with instruction-type (verbal definition, action observation, action imitation, enactment) and verb type (locomotion, object-manipulation, abstract) as within-subject factors, and standardized vocabulary scores as covariate was used to analyze the recall and comprehension data. To investigate the differences between children with good and poor language skills, separate analyses were run, in which children were split into two groups: children with a good vocabulary (A or B score on the standardized test) and children with a poor vocabulary (C, D or E score on the standardized vocabulary test).

Results

Recall test

The 4 x 3 repeated measures ANCOVA showed a significant interaction between instruction and verb type, $F(6, 264) = 2.51$, $p = 0.038$, $\eta_p^2 = .29$ with a significant effect of the covariate, $F(1, 44) = 8.15$, $p = 0.01$, $\eta_p^2 = .15$. In Figure 2 the results of the recall test are displayed for the three different verb types.

Follow-up t-tests to explore this interaction revealed no significant difference between recall of the three verb types within the verbal definition, gesture observation, and imitation conditions. In the enactment condition, however, there was a significant difference between object manipulation and locomotion verbs, $t(1, 48) = 2.57$, $p = 0.01$, $d = .47$ and a marginally significant difference between object manipulation and abstract verbs, $t(1, 48) = 1.89$, $p = 0.065$, $d = .37$, uncorrected for multiple comparisons. This finding partially confirms the hypothesis that object manipulation words are learned better than locomotion verbs and it gives slight evidence for a smaller effect of motor activation for learning abstract words than object manipulation words.

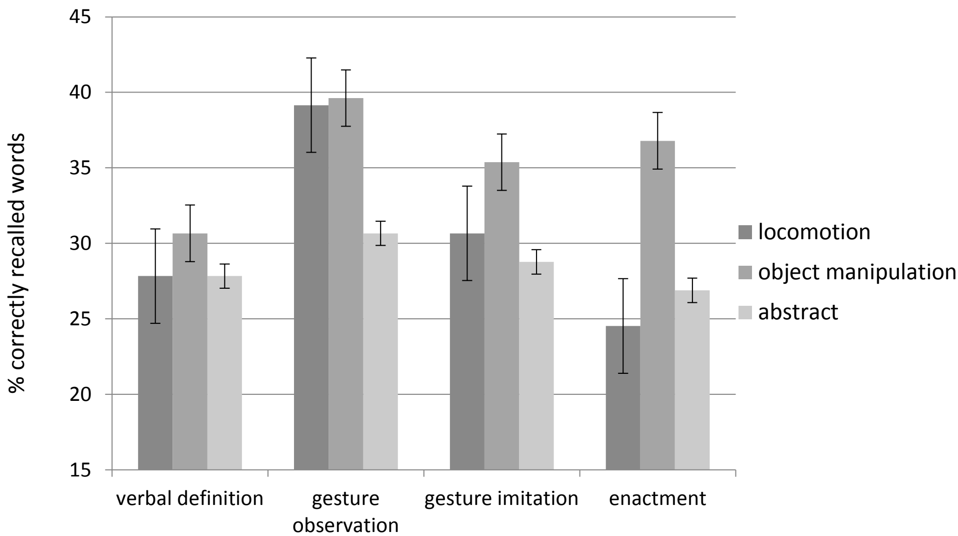


Figure 2. Recall test scores on the three verb types under the four conditions.

As for the effects of instruction within verb types, follow-up t-tests showed no significant difference for the abstract verbs in the scores on the different types of instruction. For the locomotion verbs, there was a significant increase in the number of verbs that were correctly recalled in the action observation condition compared to the verbal definition condition, $t(1, 48) = 2.1$, $p = 0.04$, $d = .39$. Furthermore, there was a marginal difference between the action observation and the action imitation condition, $t(1, 48) = 1.73$, $p = 0.09$, $d=.25$, and a significant difference between the action observation and the enactment condition $t(1, 48) = 2.50$, $p = 0.02$, $d = .53$. For the object manipulation verbs there was no significant effect of the instructional methods. These results suggest that, at least for locomotion verbs, activation of the motor system diminished the number of words learned.

To investigate whether these effects would differ for children with good and poor vocabulary scores, we performed separate t-tests in which we compared the difference in performance of the children with good and poor vocabulary skills (see Table 1). For the locomotion verbs there was no significant difference in the performance on the words learned under different conditions for these two groups. The same was true for the abstract words. However, for the object manipulation verbs a different pattern emerged. Children with good vocabulary scores, had marginally higher learning outcomes on verbs in the enactment condition than on verbs in the verbal definition condition $t(1, 20) = 1.85$, $p = 0.079$, $d = .51$, while for children with poor vocabulary scores there was no difference in performance between the four instructional conditions. For the action imitation $t(1, 42) = 2.38$, $p = 0.02$, $d = .73$, and enactment condition $t(1, 42) = 2.72$, $p = 0.01$, $d = .83$, there was a significant difference between the children with good and poor vocabulary skills. The children with good vocabulary skills performed better in the imitation and enactment conditions than children with a poor vocabulary score. Contrary to our hypothesis, children with high verbal skills benefitted from action execution while children with low verbal skills learned significantly fewer words in these conditions. To explore whether differences between children with good and poor skills were due to differences in the manner in which gestures were imitated or enacted, we looked at the videos in which the children with high and low verbal skills produced gestures. For the imitation condition, there were no obvious differences between the children with high and low verbal skills. In both groups there were children, who immediately correctly imitated the gesture they just saw, and children who produced a very small version of the gesture or children who had to be encouraged to imitate the gesture the way they had seen it (for example, a child was asked to imitate the gesture for 'to loiter' and first did this by moving her legs while sitting on a chair. Not until being prompted to imitate the gesture by actually walking did she do so). Thus, the differences in learning outcomes in the imitation condition do not seem to stem from differences in the way the gestures were imitated by children with high and low verbal skills, but are more likely to stem from a distraction from the definition that had to be learned when asked to actively produce a gesture.

Table 1. Means (and Standard Deviations) of % Items Correct on the Recall Test for Children with Good and Poor Scores on the Standardized Vocabulary Test (Good n=21, Poor n=23)

Recall test	**Instructional method**							
	Verbal definition		**Action observation**		**Action imitation**		**Enactment**	
Verb type	**Poor**	**Good**	**Poor**	**Good**	**Poor**	**Good**	**Poor**	**Good**
	M(SD)		**M(SD)**		**M(SD)**		**M(SD)**	
Locomotion	20(28)	35 (31)	35(35)	48(33)	29(29)	38(41)	26(27)	25(32)
Object manipulation	30(31)	37(32)	39(29)	46(30)	27(29)	50(34)	26(33)	54(34)
Abstract	28(29)	32(40)	29(36)	36(32)	26(37)	37(35)	24(29)	32(38)

In the enactment condition, both children from the high and low verbal skills group created gestures that were more or less similar to the ones that we created. However, in a few cases the gestures were somewhat simplified. For example, when asked to make a gesture for ‘to chisel’, which was a hitting motion towards a closed fist in our gesture, many children only made a hammering motion with one hand, and when asked to create a gesture for ‘to weed’, which in our gesture was realized by moving both hands in a wiggly manner as if moving dirt with a tool, a child would just move one hand over the ground as if moving dirt with his/her hand. Although these gestures for object-manipulation verbs were slightly different from the ones we created and were not always a fully accurate representation of the nuanced meaning of the verb, they were all meaningful. The only difference between the low and high verbal skills children that became apparent when looking at the videos was that children with low verbal skills more frequently indicated not knowing what a gesture should look like for the abstract verbs. However, there were no differences in learning outcomes between the groups on abstract verbs.

Comprehension test

The scores on the comprehension test for children with good and poor verbal skills can be found in Table 2. The 4 x 3 repeated measures ANCOVA showed a significant three-way interaction between instructional method, verb type and vocabulary score, $F(6, 252) = 2.47$, $p = 0.02$, $\eta_p^2 = .06$, with a significant effect of the covariate $F(1, 42) = 10.8$, $p = 0.002$, $\eta_p^2 = .20$.

Table 2. Means (and Standard Deviations) of % Items Correct on the Comprehension Test for Children with Good and Poor Scores on the Standardized Vocabulary Test (Good, n=21, Poor n=23)

Comprehension test	**Instructional method**							
	Verbal definition		**Action observation**		**Action imitation**		**Enactment**	
Verb type	**Poor**	**Good**	**Poor**	**Good**	**Poor**	**Good**	**Poor**	**Good**
	M(SD)		**M(SD)**		**M(SD)**		**M(SD)**	
Locomotion	20(25)	15(23)	15(24)	21(29)	9(19)	27(29)	15(24)	15(23)
Object manipulation	23(31)	36(36)	25(33)	45(35)	29(33)	40(37)	11(21)	50(42)
Abstract	11(21)	31(33)	11(21)	24(30)	14(27)	29(30)	15(23)	19(33)

There was a significant difference between object manipulation and locomotion verbs, $t(1, 48) = 2.91$, $p = 0.005$, $d = .42$ for the verbal definition condition. For the observation condition there was a significant difference between object manipulation verbs and locomotion verbs $t(1, 48) = 2.96$, $p = 0.005$, $d = .49$, and a significant difference between object manipulation and abstract verbs, $t(1, 48) = 2.58$, $p = 0.01$, $d = .52$. Also, for the action imitation condition there was again a significant difference between object manipulation and locomotion verbs, $t(1, 48) = 3.26$, $p = 0.002$, $d = .55$ and a difference between object manipulation verbs and abstract verbs $t(1, 48) = 2.98$, $p = 0.005$, $d = .50$. For the enactment condition there was a significant difference between object manipulation and locomotion verbs $t(1, 48) = 2.59$, $p = 0.01$, $d = .45$. These results are in line with the hypothesis that object manipulation verbs are learned (in this case comprehended) better than locomotion verbs and abstract verbs, which is true under all instructional conditions.

We then analyzed the differences within verb types for the different methods of instruction. For all three verb types (i.e., locomotion, object manipulation and abstract), there were no significant differences between the instructional methods. From these data it appears that gestures are not beneficial for a comprehension task. We will elaborate on this finding in the discussion.

We then performed separate analyses for the scores on the comprehension test to investigate the difference between children with high and low vocabulary scores. There were no significant differences between these two groups for the abstract words. On the locomotion verbs, the groups only differ significantly from each other in the action imitation condition $t(1, 42) = 2.51$, $p = 0.02$, $d = .77$, in which children with poor vocabulary scores performed significantly lower than children with good vocabulary scores. For the object manipulation verbs, children with good vocabularies differed significantly from children with poor vocabulary scores in the enactment condition $t(1, 42) = 3.97$, $p = 0.0003$, $d = 1.21$, while the scores do not differ significantly in any of the other conditions. In sum, children with high verbal skills benefitted more from action execution than children with low verbal skills, which was not in line with our hypothesis.

Discussion

Research has shown that gestures can foster learning in a variety of tasks, including word learning (e.g., Cook, Mitchell, & Goldin-Meadow, 2008; Tellier, 2005). However it is still unclear *why* gestures can foster learning. In this explorative study we made a first step to come closer to answering this question, by first asking what the most effective manner is in which gestures can be used in word learning (i.e., gesture observation, imitation or generation), whether gestures can promote learning of different *types* of verbs in the *first* language, whether gestures aid recall or also comprehension, and whether the effects are the same for high and low ability learners.

The hypothesis that more words would be learned in conditions in which gestures are used (hypothesis 1) was partially supported, but the effects differed depending on verb type. For locomotion verbs there was a significant improvement in the number of words learned in the gesture observation condition as compared to the verbal definition condition, a difference that was not found for the other two verb types (i.e., object manipulation and abstract verbs). This was only true for the recall test and not for the comprehension test, where we did not find any differences between instructional methods. Potential explanations for why we did not find an effect of gestures for the abstract verbs, object manipulation verbs or on the comprehension test will be discussed later on in this section. The effects of instructional condition are, however, unlikely to be due to order effects or effects of fatigue, given that the experiment was conducted over two days. If fatigue would have played a role, then we would have expected the highest scores to appear on the verbal definition condition and the imitation condition, which were the first conditions on each day. This was not the case.

For the action observation and action imitation conditions we expected that the words in the action imitation condition would be learned better than the words in the action observation condition (hypothesis 2). For the locomotion verbs, there was indeed a difference between the action observation and action execution conditions, but not in the direction we expected. Observation led to a higher recall performance than action execution (significantly higher than enactment, marginally higher than imitation). A potential explanation for this finding could be that children had to produce gestures in which they needed to walk, these gestures might have broken their concentration more than the production of other gestures for most of which children could remain seated. We further expected that verbs from the action imitation and action execution condition would be learned equally well. For all verb types (i.e., object manipulation, abstract, object manipulation verbs) there was no significant difference between words learned in the action imitation and enactment condition.

How can these findings be explained? Previous studies that found an effect on memorization when gestures were made (i.e., imitated or enacted) have mostly focussed on second language learning. In second language learning, participants already have a concept (in their first language) of the word they have to learn, while in our study the children had to create a whole new concept, making our task much more difficult. The added difficulty of simultaneously learning a word and thinking about how to imitate or enact it might have been too challenging for these children and could have caused the poorer performance in the action imitation and enactment conditions. Although this effect was only significant for the locomotion verbs, the scores for the other two verb types did not improve under these action execution conditions compared to the verbal definition condition. This working memory load explanation would also explain why particularly children with lower language ability would perform poorer when having to execute gestures; a point to which we will return later on. Furthermore, studies that found an effect of enactment (e.g., Engelkamp & Zimmer, 1997) have typically used action phrases, not a range of verb types. The material used in these prior studies was quite different from our material, which could explain why we did not find the same effect. It could also be the case that the effect of action imitation and enactment on memorization does not emerge immediately after encoding, but is only apparent after a certain delay, which is what Macedonia and Knösche (2011) found for memorization of sentences from an artificial language. Speculatively, this is also the reason why we did not find an effect of gestures on the comprehension test. Future research using a delayed test might shed light on this possibility.

In line with our hypothesis concerning the different types of verbs, the effect of gestures on learning of abstract words was smaller than the effect on locomotion and object manipulation verbs (hypothesis 3). Not only did participants score lower on recall and comprehension of abstract verbs, there was also no effect of instructional method for the abstract verbs. This finding is interesting, because it suggests that gesturing instructions may not be effective for learning words that do not have a direct link to the motor system, even though theories of embodied cognition would predict an effect of motor activation on abstract words as well.

Contrary to our hypothesis, object-manipulation verbs did not benefit more from the use of gestures during learning than did locomotion verbs (hypothesis 4), although scores were higher for object-manipulation verbs compared to locomotion and abstract verbs on the comprehension test, this effect was constant over all (non)gesture conditions. Overall, locomotion words seemed to profit from gestures more than object manipulation verbs. However, the scores on the object manipulation verbs followed the same pattern as those on the locomotion verbs, in that scores improved in the action observation condition compared to the verbal definition condition, but this improvement only reached significance for the locomotion verbs. Not finding a significant effect of gestures for the object-manipulation words was rather unexpected, because this is the verb type that has been investigated most often in studies using gestures to improve word learning. As mentioned before, however, these studies mostly used second language acquisition and the creation of a new concept might make the task too difficult. One possible reason for why we did find an effect on locomotion verbs and that these verbs benefit most from gesture observation, is that the gesture exactly portrays the meaning of the verbs. Gestures where the use of an object is depicted are cognitively more complex (Cartmill, Beilock, & Goldin-Meadow, 2012), because here some form of imagination is required, which is not the case in the relatively less complex locomotion gestures. It is possible that this benefit for locomotion verbs did not persist in the action execution conditions, because execution demands cognitive resources that might not have been available to all children, causing the disappearance of the effect (and even creating a negative effect) of executing gestures for the locomotion verbs.

Our final hypothesis concerned the potential difference in performance for the children with low verbal skills and children with high verbal skills. It predicted that children with low verbal skills would benefit more from the use of gestures during learning than would children with high verbal skills (hypothesis 5). As opposed to Frick-Horbury (2002) we found that children with high verbal skills benefitted more from imitation and enactment of gestures than children with low verbal skills. We speculatively attribute this finding to a difference in working memory capacity for the children with good and poor vocabularies, because vocabulary knowledge has been found to be related to working memory capacity (Daneman & Green, 1986). The children with good vocabularies might, therefore, have had more processing capacity available (similar to the adults in the study by Frick-Horbury, 2002), which makes them better able to execute gestures while simultaneously listening to the meaning of the words (imitation and enactment). Working memory capacity was not measured in this study, but might prove to be an informative measure in future research on this topic, because of its role in language acquisition (Ellis & Sinclair, 1996). An alternative, speculative explanation could be that children with high verbal skills make a better "link" between word and action. Thorough processing of words could lead to more grounding of the words in action. Children with poorer verbal skills could process words in a more “detached” superficial manner, causing them to benefit less from the information that gestures could provide. However, one could also argue that if children with high verbal skills already ground words in action more effectively, the gestures would not be needed to profit from motor information during learning. This explanation would therefore have to be further investigated in future research.

Lastly, the findings discussed above mostly concern the results of the recall test. On our comprehension (i.e., fill-in-the-gap) test, we did not find an effect of gestures on verb learning. Given that previous studies on the effect of gestures on word learning have mostly used free recall and recognition tests, we could not formulate a clear hypothesis about what the effect of gestures on a comprehension test would be. Our findings suggest that gestures mainly improve recall, but that there is no effect of gestures on applying newly learned information –at least on an immediate test. It could also be the case, that our test was not sensitive enough to measure comprehension. Future studies should further address this issue, by using other comprehension tests and delayed tests to see if comprehension is aided by gesturing at a delay.

An interesting question is whether the results of this study provide support for the theory of embodied cognition or can be sufficiently explained by theories regarding multisensory input (e.g., Shams & Seitz, 2008). Although both theories make similar predictions, for instance that words under a gesture observation condition would be learned better than words learned by a verbal definition only, either because of the motor activation or because of the extra modality that is used, a theory of multisensory input would have a hard time explaining why we found differential effects of gesturing for the different verb types (e.g., a beneficial effect of gesture observation for locomotion verbs but not for object-manipulation verbs). These differences can be explained when looking at the motor activation that is required for these words. The gesture for a locomotion verb is an exact reflection of the meaning of the word, while for an object manipulation word the learner has to imagine the use of a certain tool. The strength of the link to the motor system therefore differs for these verbs. From an embodied cognition approach it could, therefore, be expected that the words that have the strongest link to the motor system would benefit most from the use of gestures during learning.

In conclusion, even though this was an explorative study, it seems to provide a starting point for further research in this area, as the results of this study suggest that observing, imitating, and generating gestures may have different effects depending on verb type and language proficiency, and that effects of gesturing on learning apply only to verbs that have a direct link to the motor system. These findings should be replicated in future research, but could have important educational implications, as they show that gestures can foster learning, but not for all verb types and all children alike. Contrary to other studies we suggest that the use of gestures during learning should be viewed in a more nuanced light and that the boundary conditions for the use of gestures in instruction should be further investigated, so eventually we might also be able to formulate an answer to *why* gestures can foster learning.

Appendix

Word	Definition (English translation)	Gesture
Object manipulation		
Beitelen (to chisel)	To hit a sharp object with a hammer to carve something out.	A hitting motion towards a closed fist.
Rijten (to rip)	To tear something apart.	Moving hands away from each other with fingers pressed together as if holding paper.
Vijlen (to file)	To make something shorter or smoother with a tool made of hardened steel.	Moving both hands from left to right as if filing wood.
Schoffelen (to weed)	To plow dirt with an object.	Moving both hands in a wiggly manner as if moving dirt with a tool.
Pureren (to mash)	To mash food.	Moving one hand down with force and the other hand open as if holding a bowl.
Boetseren (to mold)	To knead something with two hands to give it shape.	Moving the fingers of both hands as if holding clay.
Dompelen (to immerse)	To keep something under water for a short amount of time.	Pressing the thumb onto the other fingers and moving both hands up and down in a flowing motion.
Zemen (to wipe)	To clean a window by wiping it.	Flat hand moving in a circular motion.
Locomotion		
Kuieren (to saunter)	To walk calmly.	A person walking slowly.
Dartelen (to frolic)	To jump around.	A person jumping in repetition.
Schrijden (to stride)	To walk with big slow steps.	A person walking with big, slow steps.
Hompelen (to hop)	To not be able to walk properly with one leg.	A person dragging one leg behind him.
Paraderen (to parade)	To walk straight in a proud manner.	A person walking, with their head held up high.
Ilsberen (to pace)	To nervously walk back and forth.	A person quickly walking back and forth.
Drentelen (to loiter)	To walk around aimlessly.	A person walking slowly in a zigzag manner.
Trippen (to trot)	To walk with short bounces.	A person walking with very small bouncy steps.

Abstract		
Kniezen (to sulk)	To be sad and worried.	A person with hand to his mouth and head down.
Afpoeieren (to dismiss)	To send someone away.	A “shoo” gesture, waving the hand in a forward manner.
Billijken (to approve)	To approve of something.	Nodding and ‘go ahead’ gesture with one hand moving forward.
Bedelen (to beg)	To ask someone for money or food.	A person holding out both hands.
Duchten (to fear)	To be afraid of something.	A person crouching with his hands in front of his face.
Ontrafelen (to unravel)	To solve something.	A person plucking at something and then holding out both hands, as if finishing pulling strings apart.
Dubben (to hesitate)	To be indecisive and not knowing what to do.	A person holding both hands out and moving these up and down as if choosing between them.
Beramen (to contemplate)	To think of something intently and at length.	A person putting the fingers beneath the chin and frowning eyebrows.

▪ Chapter 3

Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance

This chapter has been published as:

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica, 144,* 173-179. doi: 10.1016/j.actpsy.2013.05.013.

Abstract

Research has shown that observing and imitating gestures can foster word learning and that imitation might be more beneficial than observation, which is in line with theories of Embodied Cognition. This study investigated when imitation of gestures is most effective, using a 2 x 2 x 2 x 3 mixed design with between-subjects factors Imitation during Encoding (IE; Yes/No) and Imitation during Retrieval (IR; Yes/No), and within-subjects factors Time of Testing (Immediate/Delayed) and Verb Type (Object manipulation/Locomotion/Abstract). Primary school children (*N* = 115) learned 15 novel verbs (five of each type). They were provided with a verbal definition and a video of the gesture. Depending on assigned condition, they additionally received no imitation instructions, instructions to imitate the gesture immediately (i.e., during encoding; IE), instructions to imitate (from memory) during the first posttest (i.e., during retrieval; IR), or both (IE-IR). Based on the literature, all three imitation conditions could be predicted to be more effective than no imitation. On an immediate and delayed posttest, only the object-manipulation verbs were differentially affected by instructional method, with IE and IR being more effective than no imitation on the immediate test; IE-IR and no imitation did not differ significantly. After a one week delay, only IR was more effective than no imitation, suggesting that imitation during retrieval is most effective for learning object-manipulation words.

Introduction

Gestures, defined here in accordance with the Oxford English Dictionary as “movements of the body or parts of the body”, are ubiquitous and serve an important function in communication (Goldin-Meadow, 1999). However, because gestures are even used when they have no obvious function (e.g., when there is no visual contact with the interlocutor; Rimé, 1983), they are thought to have not only communicative functions, but cognitive ones as well (Kita, 2000). Given that language and the motor system are closely connected (Fischer & Zwaan, 2008) and that speech and gestures even seem to share the same communication system (Bernardis & Gentilucci, 2006), theories of Embodied Cognition would predict that gestures have a function in language learning.

In line with this assumption, it has been shown that *observation* of gestures can improve word learning. A study in which five year old children received verbal instructions with or without observation of congruent gestures, showed that observation improved memorization of a list of concrete nouns and descriptive adjectives (Tellier, 2005). In addition, gesture observation seems to have benefits for application of new information as well (Ping & Goldin-Meadow, 2008).

Next to observation, gesture *imitation* has been shown to improve word learning. For example, children can remember more items from a list with familiar concrete nouns and descriptive adjectives when imitating gestures that accompany the words than when only observing these gestures, an effect that was found both for words from the first language (Tellier, 2005) as for words from a second language (Tellier, 2008). These findings suggest that imitation of gestures might have a beneficial effect on learning the meaning of verbs over and above mere observation of the same gestures. However, these studies applied imitation during encoding, immediately after observation. An interesting open question that we address here is what the effects on test performance would be of imitating the gestures that were observed during encoding, in the retrieval phase. Whereas imitation while hearing new words may facilitate encoding of information, imitation during retrieval (or test taking) might provide children with additional cues that could foster the retrieval of previously learned words from memory. Imitation at both of these stages might be even more effective for learning, aiding both encoding and retrieval.

Imitating gestures during encoding

In the speech production literature, it has been hypothesized that gestures can facilitate speech production, because they facilitate conceptual storage before the information has a linguistic form. That is, gestures would translate spatio-motoric knowledge into linguistic units (Kita, 2000). A similar mechanism might apply to encoding new verbal information. The Image Maintenance Theory (De Ruiter, 1998; Wesp, Hesse, Keutmann, & Wheaton, 2001) proposes that gestures facilitate the maintenance of spatial representations in working memory. According to this theory, gestures activate images in working memory and can reactivate visuo-spatial information which might facilitate encoding. Indeed, as mentioned above, imitation during encoding has been found effective in language learning (e.g., Tellier, 2007, 2008; see also findings with adults by Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011).

Imitating gestures during retrieval

From domains other than language learning, there are indications that spontaneous gesturing (i.e., not specifically imitating previously observed gestures) during the recollection of a previous event, improves recall compared to when children are prohibited from gesturing (Stevanoni & Salmon, 2005). Similarly, recall of verbal material has been shown to improve when adults are cued with self-generated hand gestures compared to when they are not cued (Frick-Horbury, 2002). Moreover, when people were prohibited from gesturing during speech with spatial content, they were less fluent than when gesturing was allowed, which shows that lexical access is disrupted (Rauscher, Krauss, & Chen, 1996). These findings are in line with the Lexical Retrieval Hypothesis (Rauscher et al., 1996), which states that gestures facilitate retrieval of items from the mental lexicon. So even though none of these studies looked specifically at imitating previously observed gestures, these findings do suggest that imitation of a gesture during retrieval could facilitate the recall of previously learned words.

Imitating gestures during encoding and retrieval

Finally, the combination of imitation during encoding as well as retrieval might be most effective for word learning, according to the Encoding Specificity Principle (Tulving & Osler, 1968). This hypothesis states that retrieval of information is more likely when the overlap between the context at encoding and the context at retrieval is high. For example, when participants were presented with weakly associated cue-target word pairs, recall of the target words was higher when participants were cued with the same weakly associated word, than when they were cued with another weakly associated cue or when free recall was used (Tulving & Osler, 1968). Also, when memorizing a list of words under water, recall of the words was better when words were retrieved under water, than when retrieval took place on dry land (Godden & Baddeley, 1975) showing that the context in which a word is learned, can aid recall. Although these studies did not use gestures, the Encoding-Specificity Hypothesis suggests that an effective way to improve recollection would be to imitate observed gestures both during the encoding and the retrieval phase. The gestures made during encoding are then part of the mental representation of the new word, and activating this part of the mental representation during retrieval might facilitate recall.

The present study investigates the effectiveness of these three different ways of imitating observed gestures, on 9-to-10 year old children's learning of novel verbs in their first language.

Effects of imitation on verb learning

Verbs are a particularly interesting word class for studying the effects of imitating gestures on learning, given the existence of verbs that have a direct link to the motor system (e.g., to stride, to chisel) and verbs that do not have a direct link to the motor system (e.g., to dismiss, to recruit). Most studies that have investigated the role of the motor system in understanding language using verbs, have used object manipulation verbs or functionally manipulable verbs (e.g., Rueschemeyer, Lindemann, Van Rooij, Van Dam, & Bekkering, 2010; Tessari, Canessa, Ukmar, & Rumiati, 2007; Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). However, object-manipulation verbs are just one class of action verbs and it is important to investigate whether effects of imitating gestures during encoding, retrieval, or both, are the same for various types of verbs, such as locomotion verbs (e.g., to stride) or abstract verbs (e.g. to dismiss).

Whereas object-manipulation verbs are mostly manual verbs, locomotion verbs require leg activation. These verbs therefore activate different parts of the motor cortex (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Tettamanti et al., 2005, see also: Hickok, 2010). Because of the close coupling between language and manual gestures (Gentilucci & Corballis, 2006), it is possible that verbs coupled to manual gestures might be learned better through imitation than verbs coupled to non-manual gestures, as is the case for the locomotion verbs. Furthermore, while both object-manipulation and locomotion verbs can be goal-directed, locomotion verbs do not always denote intentional goal-directed activities (e.g., to stroll). Research has shown that the mirror neuron system (Rizzolatti & Craighero, 2004) is more active for goal-directed actions than for non-goal-directed actions (Nyström, Ljunghammar, Rosander, & von Hofsten, 2011), which might aid imitation. Imitation involves a decomposition of motor patterns into constituent components and later a reconstruction of the action pattern from these components. This decomposition is guided by an interpretation of the motor pattern as goal directed behavior (e.g., Bekkering, Wohlschläger, & Gattis, 2000), suggesting that actions may be harder to imitate when they are not goal-directed. Thus, these findings suggest that motor activation resulting from imitating gestures could be stronger and therefore perhaps more beneficial for learning object-manipulation verbs than for learning locomotion verbs.

Both object manipulation and locomotion verbs are action verbs and have an obvious connection to the motor system, which abstract verbs do not have. According to pure theories of embodied cognition, cognition is completely grounded in the sensory-motor system, and accordingly, this view would predict that such grounding should exist for abstract concepts as well (e.g., Pecher, Boot, & van Dantzig, 2011; Wilson, 2002), in which case imitation can be expected to facilitate learning of abstract verbs as well. For instance, given that the abstract concept of time can be understood in terms of spatial experience (Boroditsky & Ramscar, 2002), abstract concepts could refer indirectly to bodily experiences (Borghi, Caramelli, & Setti, 2005).

The present study

In summary, we will investigate the effects of different instructional conditions (i.e., no imitation, imitation during encoding, imitation during retrieval, imitation during both encoding and retrieval) on learning of novel verbs in the children's first language (Dutch). Learning will be measured both immediately and at a delay. This is done not only because it has been shown that the effect of observing and producing gestures might not emerge immediately, but only after a certain delay (e.g., Cook & Fenn, 2010; Macedonia & Knösche, 2011), but also because it is relevant from a practical (i.e. educational) perspective to investigate whether effects of imitating gestures persist or change over time. For word learning it is not that relevant to be able to remember a definition straight after you have heard it. It is far more important to remember what the word means after a longer interval, so that the knowledge of the word can actually be used when it is encountered in other contexts and the word can be added to the active vocabulary.

Our review of the literature leads to several hypotheses concerning the effect of the different instructional conditions in relation to verb type (i.e., locomotion, object manipulation, and abstract). We expect children to learn more words in the imitation conditions (during encoding, during retrieval, or both) than in the no imitation (i.e., observation only) condition, based on studies that found an effect of gesture imitation over gesture observation (e.g., Tellier, 2007). We expect this effect to be largest for the object manipulation verbs, because of their direct link with the motor system, their goal-directedness, and the fact that they are linked to manual gestures.

Concerning the moment at which gestures can best be imitated, there is evidence from different theoretical perspectives suggesting that imitating gestures during encoding (Image Maintenance Theory), retrieval (Lexical Retrieval Hypothesis), or both (Encoding Specificity Hypothesis), could all have beneficial effects on learning new words.

As for immediate vs. delayed test performance, it can be hypothesized that delayed performance will be lower than immediate performance, but regarding effects of instructions no clear hypothesis can be formulated based on the literature concerning the effect of the different gesture imitation conditions after a one week delay, so this is an open question explored here.

Method

Participants and design

Participants were 120 Dutch primary school children (61 male, 59 female), between nine and eleven years of age (*M* = 10, *SD* = 0.6), who were recruited from three primary schools in the same area of the Netherlands. Children, who were not born in the Netherlands, were excluded from data analysis. This information was obtained from the school teachers after data collection and resulted in the exclusion of five children (2 from the imitation during retrieval and 3 from the imitation during both encoding and retrieval condition), so the analyses are based on 115 participants. We tested primary school children of approximately ten years old because these children are still in the process of acquiring new words in their first language, while being old enough to be able to participate in a study like this one, which requires sustaining attention for the duration of the study, understanding the instructions, and learning a substantial number of new words within a relatively short time.

A 2 x 2 x 2 x 3 mixed design was used, with between-subjects factors Imitation during Encoding (IE; Yes/No) and Imitation during Retrieval (IR; Yes/No), and within-subjects factors Time of Testing (Immediate/Delayed) and Verb type (Object manipulation /Locomotion / Abstract) (Fig. 1). Participants were pseudo-randomly assigned to one of the four between-subjects conditions, matching for language ability of which an index (children's performance on a standardized test) was obtained from the teachers: No Imitation (*n* = 31), Imitation during Encoding (IE; *n* = 30), Imitation (from memory) during Retrieval on the first posttest (IR; *n* = 27), or both (IE-IR; *n* = 27).

Materials

Verbs & Gestures. The stimuli consisted of 15 Dutch verbs in total, representing three different verb types (five of each): locomotion (e.g., to stride), object-manipulation (e.g., to chisel) or abstract (e.g., to contemplate). The verbs from these different categories were matched for phoneme length and for lexical frequency by determining the mean logarithmic lemma frequency per million for each verb type using the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). To minimize the chances that children were familiar with the Dutch verbs that they were taught during the experiment, the Measure of Lexical Richness (MLR) was used. The MLR is a database derived from words that children are most likely to come into contact with at primary school. These words are divided over nine word lists (Schrooten & Vermeer, 1994), which are ordered in terms of lexical frequency. List one contains the highest frequency words, which children acquire relatively early, while list nine contains the lowest frequency words, which are learned in the upper grades of primary school. The verbs in this experiment only appeared on word lists eight or nine, or they did not appear on these lists at all, meaning that these words are most likely not learned before age twelve. This method minimized the chance that the children knew the verbs prior to the experiment. All material was pilot-tested for this age group before the experiment, to ascertain that the complexity of the verbs and the degree of difficulty of the posttest was appropriate for the age group and that the gestures that were shown in the videos, gave an accurate representation of the meaning of the verbs.

For each verb a congruent gesture was created. This gesture was made by a young adult actor. Our choice for using an adult model was motivated by considerations of ecological validity and practical implications, meaning that in the classroom, gestures during instructions will typically be made by a teacher. All actions were depicted as realistically as possible and might therefore be described as pantomimes. The object-manipulation verbs were manual, the locomotion verbs involved leg movement (e.g., seeing someone striding), and for the abstract verbs both hands and legs were used. For example, for the abstract verb ‘to dismiss’ (defined as ‘to send someone away’), the children saw a person making a “shoo” gesture, waving the hand in a forward manner, while for ‘to fear’ (defined as ‘to be afraid of something.'

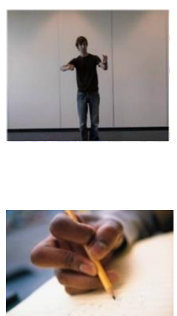
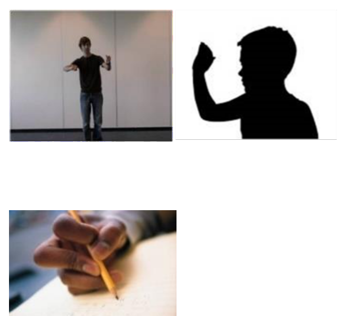
No imitation	Imitating during Encoding	Imitating during Retrieval	Imitating during Encoding + Retrieval
Verbal definition 	Verbal definition 	Verbal definition 	Verbal definition 
Verbal definition + observation of a congruent gesture 	Verbal definition + observation and immediate imitation of a congruent gesture 	Verbal definition + observation and imitating the congruent gesture during retrieval 	Verbal definition +observation and immediate imitation and imitation during retrieval of a congruent gesture 

Figure 1. Illustration of the experimental conditions, showing the separate elements of each condition.

NB: this word seems easy in its English translation but was a more complex word in Dutch not yet known to the children at this age), they saw a crouching person with his hands in front of his face.

Instructional conditions. In the *no imitation condition* the children were offered a verbal definition for a novel verb twice. The child was asked to listen to the definition and try to remember it. The first presentation consisted of a digital video of a young adult female providing the verbal definition standing completely still. The second presentation consisted of a video of the young adult male actor making a gesture congruent with the meaning of the novel verb, with the verbal definition provided as a voice-over. In the *imitation during encoding condition,* children were presented with the same materials, but were required to immediately imitate the gestures they observed. In the *imitation during retrieval condition* the children were presented with the same materials and instructions as the no imitation condition in the learning phase, but they were asked to imitate the gesture that they had been previously presented with during the immediate recall test. Lastly, in the *imitation during encoding and retrieval condition* children were presented with the same materials and instructions to imitate the gesture both during the learning phase and during the immediate recall test.

Posttest. The posttest consisted of two parts. The first part was a recall test, in which children were asked to provide a definition of the verbs. This was a paper-and-pencil test, but on the immediate test, the target words were read orally to the children and children answered orally before writing the answer down to prevent any difficulties with decoding a written word from interfering with the test performance of the child. However, in the delayed posttest children read the sentences themselves and wrote down their answers without first answering orally, because this test was administered in a group session. Because a recall test does not measure comprehension of the verb, the second part of the posttest consisted of a fill-in-the-gap test. This test contained one sentence for each of the previously learned verbs, which resulted in five sentences per condition. No list of possible answers was given to prevent the test from turning into a recognition task. An example of a sentence from the fill-in-the-gap test was: “The carpenter wanted to build a closet and took a piece of wood and started to….

(chisel)." All correct answers were infinitives, given that conjugations would add a degree of difficulty.

Procedure

Before the experiment, scores on a standardized vocabulary test were collected from the schools for all children participating in the study. Scores on this vocabulary test consist of five categories, which have cut-off scores based on a national norm group: A (i.e., score in the top 25% of scores nationally), B (i.e., score in the next 25% range, just above the national average), C (i.e., score in the next 25%, just below the national average), D (i.e., score in the next 15%, scoring well below the national average), E (i.e., score in the 10% lowest scores).

In individual sessions of approximately 20 min. duration, the children learned 15 verbs in three blocks of five under imitation instructions depending on their assigned condition. The 15 verbs consisted of five verbs of each verb type (i.e., locomotion, object manipulation, abstract). After each block of five verbs, which consisted of a mix of the different verb types, the children were presented with the posttest for those verbs, with imitation instructions depending on their assigned condition, after which they continued to learn the next set of five verbs, were posttested, learned the last set of five verbs and were posttested on those five verbs. Verbs were counterbalanced across lists, to avoid order effects. Approximately one week later, the children were presented with the same posttest for all 15 verbs at once in a group session and without any condition-specific instructions, to investigate how well the knowledge was retained. Children filled out the posttest at their own pace and could go back to a previous exercise.

Scoring

Both the recall test and the fill-in-the-gap test were scored by awarding one point for each correct answer. In the recall task children could also be rewarded a half point when one element of the definition was missing (e.g., 'to hit a sharp object' instead of 'to hit a sharp object with a hammer' for 'to chisel'). Ten percent of the data was coded by two coders, and the second coder was blind to the experimental condition of each response. The interrater reliability was Kappa = .80 for the recall test and Kappa = .98 for the comprehension test.

Results

Recall test

Data were analyzed with a repeated measures ANCOVA with Time of Testing (Immediate vs. Delayed) and Verb Type (locomotion vs. object manipulation vs. abstract) as within-subjects factors, and Imitation during Encoding (Yes vs. No) and Imitation during Retrieval (Yes vs. No) as between-subject factors. Scores on the standardized vocabulary test were used as a covariate. Mean scores and standard deviations for both tests are given in Table 1. A complete overview of main and interaction effects can be found in Table 2. The ANCOVA showed a main effect of Time of Testing, $F(1, 110) = 17.39$, $p < .001$, $\eta_p^2 = .14$, and a significant effect of the covariate, $F(1, 110) = 64.69$, $p < .001$, $\eta_p^2 = .37$. As expected, the scores on the delayed test were significantly lower than on the immediate test for all verb types.

Table 1. Means (and Standard Deviations) of % Items Correct on the Immediate (Imt.) and Delayed (Del.) Recall Test

	Condition							
	No Imitation		**IE**		**IR**		**IE-IR**	
Verb type	Imt.	Del.	Imt.	Del.	Imt.	Del.	Imt.	Del.
Locomotion	59(28)	37(25)	64(26)	47(33)	58(23)	38(24)	61(25)	45(25)
Object manipulation	27(24)	16(19)	38(27)	25(30)	43(25)	29(26)	37(20)	23(25)
Abstract	50(24)	28(26)	55(30)	35(29)	49(24)	30(25)	51(25)	34(26)
Overall Average	**45(25)**	**27(23)**	**52(28)**	**36(31)**	**50(24)**	**32(25)**	**50(23)**	**34(25)**

Table 2. Overview of Main and Interaction Effects (significant results are printed in bold)

Overall analyses	**Recall test**	**Comprehension test**
Main effects		
Verb	$\mathbf{\textit{F}(2, 109) = 6.41, \textit{p} = .002, \eta_p^2 = .06}$	$F(2, 109) = .23, p = .79, \eta_p^2 = .002$
Time	$\mathbf{\textit{F}(1, 110) = 17.39, \textit{p} < .01, \eta_p^2 = .14}$	$F(1, 110) = .04, p = .85, \eta_p^2 < .001$
Interaction effects		
Verb * Imitation during encoding	$F(2, 109) = .72, p = .49, \eta_p^2 = .01$	$F(2, 109) = .02, p = .98, \eta_p^2 < .001$
Verb * Imitation during retrieval	$F(2, 109) = 2.59, p = .08, \eta_p^2 = .02$	$F(2, 109) = .78, p = .46, \eta_p^2 = .01$
Verb * Imitation during encoding * Imitation during retrieval	$F(2, 109) = 2.30, p = .10, \eta_p^2 = .02$	$F(2, 109) = 1.68, p = .19, \eta_p^2 = .02$
Verb * Vocabulary score	$F(2, 109) = 1.21, p = .30, \eta_p^2 = .01$	$F(2, 109) = .90, p = .41, \eta_p^2 = .01$
Time * Verb	$F(2, 109) = .36, p = .69, \eta_p^2 = .03$	$F(2, 109) = 1.19, p = .31, \eta_p^2 = .01$
Time * Verb * Vocabulary score	$F(2, 109) = .04, p = .96, \eta_p^2 < .001$	$F(2, 109) = .21\ p = .81, \eta_p^2 = .002$
Time * Verb * Imitation during encoding	$F(2, 109) = .56, p = .57, \eta_p^2 = .005$	$F(2, 109) = 1.98, p = .14, \eta_p^2 = .02$
Time * Verb * imitation during retrieval	$F(2, 109) = .41, p = .63, \eta_p^2 = .004$	$F(2, 109) = .13, p = .89, \eta_p^2 = .001$
Time * Verb * Imitation during encoding * Imitation during retrieval	$F(2, 109) = .02, p = .98, \eta_p^2 < .001$	$F(2, 109) = .05, p = .95, \eta_p^2 < .001$
Between subject effects		
Vocabulary score	$\mathbf{\textit{F}(1, 110) = 64.69, \textit{p} < .001, \eta_p^2 = .37}$	$\mathbf{\textit{F}(1, 110) = 95.80, \textit{p} < .001, \eta_p^2 = .47}$
Imitation during encoding	$F(1, 110) = 1.71, p = .17, \eta_p^2 = .02$	$F(1, 110) = .35, p = .56, \eta_p^2 = .003$
Imitation during retrieval	$F(1, 110) = .15, p = .70, \eta_p^2 = .001$	$F(1, 110) = .13, p = .72, \eta_p^2 = .001$
Imitation during encoding * Imitation during retrieval	$F(1, 110) = 3.00, p = .08, \eta_p^2 = .03$	$F(1, 110) = .65, p = .11, \eta_p^2 = .02$

Separate follow-up analyses for different verb types for the recall test		
Imitation during encoding * Imitation during retrieval	Locomotion	$F(1, 110) = .67, p = .41, \eta_p^2 = .06$
	Abstract	$F(1, 110) = .48, p = .49, \eta_p^2 = .004$
	Object manipulation	$\mathbf{\textit{F}(1, 110) = 8.16, \textit{p} = .005, \eta_p^2 = .07}$

There was also a significant main effect of Verb Type, $F(2, 109) = 6.41$, $p = .002$, $\eta_p^2 = .06$. The interaction between the between-subject factors Imitation during Encoding and Imitation during Retrieval, $F(1, 110) = 3.00$, $p = .08$, $\eta_p^2 = .03$, and the interaction between Verb Type, Imitation during Encoding, and Imitation during Retrieval, $F(2, 109) = 2.30$, $p = .10$, $\eta_p^2 = .02$ were not significant. This could be partly due to the fact that the verbs differed on various dimensions that we were unable to control for. For this reason, we explored the effects for the verbs separately, by pursuing follow-up analyses on each of the verb types in an exploratory fashion. These analyses showed that there was only a significant interaction between Imitation during Encoding and Imitation during Retrieval on the object-manipulation verbs, $F(1, 110) = 8.15$, $p = .005$, $\eta_p^2 = .07$, but not on the locomotion and abstract verbs (see Table 2). The interaction effect for the object manipulation verbs is depicted in Figure 2.

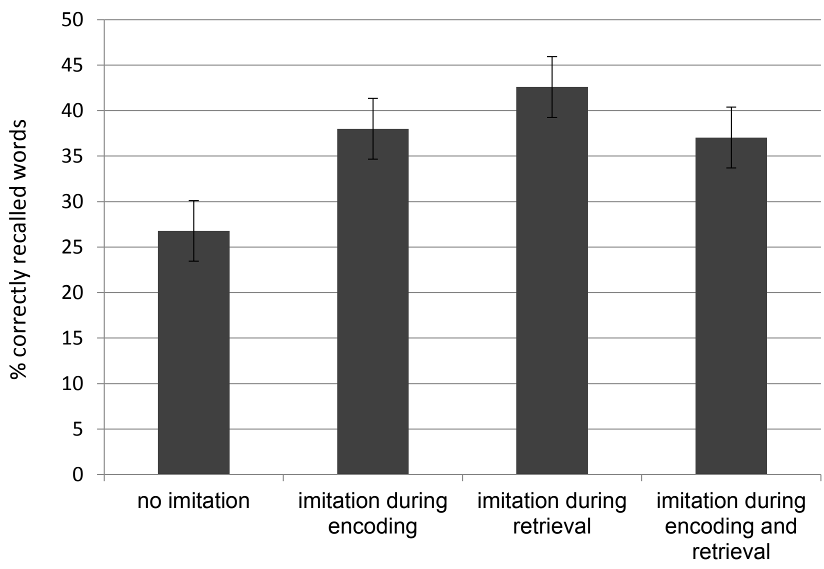


Figure 2. Recall test scores on the object manipulation verbs under the four different conditions.

This interaction on object-manipulation verbs was tested further by analyzing whether each of the imitation conditions (i.e., IE, IR, IE-IR) differed significantly from the No Imitation condition, using t-tests with Bonferroni adjusted alpha levels of 0.017 per test (.05/3). There was a significant difference between the No Imitation and IR condition in recall performance on both the Immediate test, $t(57) = 3.28$, $p = .001$, $d = .63$, and the Delayed test, $t(57) = 2.53$, $p = .013$, $d = .56$. There was also a significant difference between the No Imitation and IE condition in recall performance on the Immediate test, $t(60) = 2.44$, $p = .016$, $d = .42$, but not on the Delayed test, $t(60) = 1.84$, $p = .07$, $d = .34$, and there was no significant difference between the No Imitation condition and the IE-IR condition in recall performance on either the Immediate, $t(57) = .78$, $p = .07$, $d = .41$, or Delayed test ($t(57) = 1.01$, $p = .31$, $d = .30$).

Comprehension test

In Table 3 the scores on the immediate and delayed comprehension test are given. A complete overview of main and interaction effects can be found in Table 2. There was only a significant effect of the covariate vocabulary score, $F(1, 110) = 95.80$, $p <.001$, $\eta_p^2 = .47$.

Table 3. Means (and Standard Deviations) of % Items Correct on the Immediate (Imt.) and Delayed (Del.) Comprehension test

	Condition							
	No Imitation		**IE**		**IR**		**IE-IR**	
Verb type	Imt.	Del.	Imt.	Del.	Imt.	Del.	Imt.	De.
Locomotion	42(27)	38(27)	40(25)	48(34)	39(25)	39(29)	41(26)	41(29)
Object manipulation	24(23)	33(29)	32(29)	43(38)	30(28)	44(33)	32(31)	33(31)
Abstract	32(27)	41(33)	31(32)	49(35)	26(23)	44(35)	31(27)	41(32)
Overall Average	**33(26)**	**37(30)**	**34(29)**	**47(36)**	**32(25)**	**42(32)**	**35(28)**	**38(31)**

Discussion

We investigated whether imitation and especially the moment of imitation of observed gestures would affect children's test performance on novel verbs immediately after studying and after a one week delay. The hypothesis that gesture imitation would lead to better recall than mere gesture observation was only partially supported, as the exploratory follow-up analyses showed that an effect of imitation was only present for the object-manipulation verbs and only for imitation during encoding or retrieval, but not both. There was no effect of imitation for the locomotion and abstract verbs. There was an effect of test moment, with recall scores on the delayed test being lower, and we also found a main effect of verb type, but this cannot be meaningfully interpreted because despite matching for length and frequency, some types of words might have been easier to learn than others.

As to why imitation was not effective for abstract and locomotion verbs, we can only speculate. For the abstract verbs it might be the case that observing or imitating gestures is not beneficial for learning, because these verbs do not have a direct link to the motor system. That is, our results regarding abstract verbs do not seem to be compatible with pure embodied accounts or wide-scope embodied standpoints that claim that cognition does not need any amodal representations and that cognition is completely grounded in the sensory-motor system (e.g., Pecher, Boot, & van Dantzig, 2011; Wilson, 2002). Instead we only found effects of gesture imitation for one type of concrete verb that has a direct link to the motor system (i.e., object-manipulation verbs), and this finding is more in line with less strict or intermediate theories of embodied cognition (theories in between pure embodied and pure disembodied), which suggest that cognition is based on both sensory-motor and amodal processing. An example of such a theory is the 'grounding by interaction framework' (Mahon & Caramazza, 2008), which states that sensory-motor information can enrich amodal information and create deeper conceptual processing. Furthermore, within this view, sensory-motor grounding is not necessary for abstract processing, which is congruent with our findings.

Although locomotion verbs do have the direct link to the motor system that abstract verbs do not have, there was no effect of imitating gestures on learning of these words. One possible explanation for the lack of effect of imitation on locomotion verb learning could be that these involved non-manual gestures. There is an especially strong link between language and manual gestures (Gentilucci & Corballis, 2006), and therefore imitating manual gestures might have more effect on language learning than imitating non-manual gestures, such as those that were coupled to the locomotion verbs. Another explanation for why the effects of imitating gestures were only present for the object-manipulation verbs might be that these verbs are goal-directed. That is, imitation involves a decomposition of motor patterns into constituent components and later a reconstruction of the action pattern from these components. This decomposition is guided by an interpretation of the motor pattern as goal-directed behavior (e.g., Bekkering et al., 2000), suggesting that actions may be harder to imitate when they are not goal-directed. It should be noted that both explanations are only tentative, and should be investigated in future research.

In sum, it seems that the effects of the instructional methods on learning only differed for object-manipulation verbs. For the recall test it seems that both gesturing during encoding and gesturing during retrieval were effective for immediate recall, which is in line with our predictions based on the Image Maintenance Theory and the Lexical Retrieval Hypothesis, respectively. However, a combination of the two was not more effective than mere gesture observation, although the data presented in Table 1 suggest it was not worse than the other imitation conditions either. This was somewhat surprising; however, given that there is no prior research on this issue, we cannot provide an explanation for this finding. On the delayed test, only the condition that engaged in imitation during retrieval on the immediate test, outperformed the no imitation condition.

For the comprehension test, the results were quite different. It appears that the ability to apply the new words in the context of a sentence was not particularly improved by the imitation of gestures either during encoding or during retrieval. Interestingly, as opposed to the results on the recall test, where scores were significantly lower on the delayed test than on the immediate test, for the comprehension test there was no main effect of time. The comprehension task might have been easier after one week for a practical reason. That is, in the immediate test children could not see the words they had to fill out in the comprehension test, while during the delayed test they could see these words. Therefore, it might have been easier for the children to couple the words to the right sentence context. However, despite these practical differences between the two times of testing, the fact that children's performance did not worsen on the comprehension test suggests that when seeing the words, children might not remember the precise meaning, but are able to couple the right word with the right sentence, for which a global mental representation of the word's meaning is necessary. In a future study both immediate and delayed posttest could be administered in individual sessions to remove the practical differences between the test moments.

Concerning the strength of the instructional conditions, it should be noted that we had an adult perform the gestures in the videos that the children observed. In general, it is not entirely clear whether peer models are more effective for observational learning than adult models. Peer-models do seem to have a positive influence on self-efficacy compared to adult models (Schunk & Hanson, 1985), but the effects of peer and adult modeling on learning performance show mixed results (Schunk, 1987). However, specifically concerning modeling of gestures, it has been suggested that the effects of motor resonance might be reduced when children are shown adults' gestures rather than other children's gestures (e.g., Liuzza, Setti, & Borghi, 2012). Even though children are very much used to seeing adult teachers gesture during instruction, it is possible that gestures made by another child might have been easier to imitate.

Lastly, we had no clear hypothesis on the effect of immediate or delayed recall of newly learned words in combination with gesture observation or imitation. In this study only one effect of gesture imitation that was found on the immediate test, remained on the delayed test (i.e., the significant improvement in the number of words learned in the retrieval condition compared to the observation condition). Contrary to the findings by Macedonia and Knösche (2011), we did not find any effects after a delay that were not present at the immediate recall test.

In conclusion, this study tentatively suggests that imitation of gestures either during learning or during an immediate recall test can have an effect on the number of verbs that is correctly recalled on both an immediate recall test and a delayed test after one week, and that the time at which the gestures are imitated (i.e., during encoding, retrieval, or both) affects recall performance. These effects seem, however, dependent on verb type, given that they only occurred for object-manipulation verbs, while locomotion and abstract verbs were unaffected by instructional method. Imitation could potentially be a useful tool in word learning, but the conditions under which it can be successfully used should be further investigated.

▪ Chapter 4

The effects of gesture observation versus imitation on children's and adult's learning of novel and known action verbs

This chapter has been submitted for publication as:

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *The effects of gesture observation versus imitation on children's and adult's learning of novel and known action verbs.* Manuscript submitted for publication.

Abstract

Word learning is an important part of the acquisition of a new language. It has been suggested that sensorimotor activation evoked by gesturing could improve word learning. It is, however, not entirely clear yet whether the effect becomes stronger when students not only observe an instructor's gestures but also imitate those gestures, and whether the effects are similar for children and adults. Therefore, we directly compared the effects of gesture observation versus imitation on action verb learning in children and adults. In addition we examined a factor that might play a role in the effectiveness of gestures for word learning, namely the novelty of the word of which the definition has to be learned. Results showed that adults outperformed children and performance was higher for known words, but word familiarity did not interact with gesture condition. There was a significant age by gesture condition interaction which suggested that for children, observation and imitation were equally effective, whereas for adults, observation was more effective than imitation, although this follow-up analysis was not statistically significant. This study, therefore, suggests that gestures may have differential effects on learning words in the first language for children and adults, which may be related to developmental differences in working memory capacity.

Introduction

When we talk, we gesture; regardless of whether we are talking to someone face-to-face or on the phone (e.g., Rimé, 1983). Even blind people gesture to each other (Iverson & Goldin-Meadow, 1998). Such findings, showing that we gesture even in absence of communicative purposes, have led to the idea that gestures serve not just communicative, but also cognitive purposes (e.g., Kita, 2000; Pouw, De Nooijer, Van Gog, Zwaan, & Paas, 2014; Wesp, Hesse, Keutmann, & Wheaton, 2001). That assumption is further supported by findings that gesturing (both observing an instructor's gestures or making gestures oneself) can improve learning, for example, math learning (e.g., Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Cook, Yip, & Goldin-Meadow, 2012; Goldin-Meadow, Cook, & Mitchell, 2009), problem solving (e.g., Alibali, Spencer, Knox, & Kita, 2011; Chu & Kita, 2011; Garber & Goldin-Meadow, 2002; Thomas & Lleras, 2009), or language learning, which is the focus of the present study (e.g., De Nooijer, Van Gog, Paas, & Zwaan, 2013; Tellier, 2005).

Why gestures can support learning, including language learning, is explained via theories of embodied cognition. According to these theories, a mental simulation is created when we read or hear a word. Mental simulations are "the reenactment of perceptual, motor, and introspective states acquired during interaction with the world, body, and mind'' (Barsalou, 2008, p.618). This mental simulation results in pre-motor activity (e.g., Willems, Toni, Hagoort, & Casasanto, 2010). Thus, language and action are closely connected (e.g., Fischer & Zwaan, 2008) and can influence each other. This is illustrated, for instance, by findings showing that speakers who are not allowed to gesture while verbally expressing spatial information are less fluent in their speech than when co-speech gestures are made (e.g., Morsella & Krauss, 2004; Rauscher, Krauss, & Chen, 1996). Moreover, it has been shown that when merely hearing a word like 'to kick', areas of the brain's motor system are activated (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Tettamanti et al., 2005).

From an embodied view on word learning it would, therefore, be expected that the use of gestures during learning would lead to a stronger activation of the motor system (in the form of mental simulations). This might improve word learning by reducing working memory load and / or enriching the mental representation of the word (e.g., Goldin-Meadow, 2010).

With regard to working memory load, the cognitive theory of multimedia learning (e.g., Mayer, 2014) suggests that working memory load is reduced and learning is enhanced when information is presented in different modalities because information processing is then distributed among different working memory channels (Baddeley, 1999), thus making more optimal use of limited working memory capacity. At the same time, because word meanings are in part made up of motor information, using gestures during learning that are congruent with what has to be learned, could lead to the creation of a deeper understanding and better consolidation of the meaning of the new word. These two explanations for why gestures can support (language) learning are not necessarily mutually exclusive (e.g., Goldin-Meadow, 2010). However, the use of gestures during word learning can be implemented in different ways, and as will become clear from the literature review below, it is still unclear what the most effective way is, and what other factors (e.g., individual or word characteristics) could play a role in the effectiveness of gestures on word learning.

Gesture observation during word learning

When learning a new language (whether it is the first or second language), acquiring a substantial vocabulary is crucial for success (e.g., Folse, 2004), so it is not surprising that many studies have investigated what techniques are effective for improving vocabulary learning. For instance, it has been investigated whether vocabulary learning could be improved by means of adding context in the form of concept wheels (Vacca & Vacca, 2002), combining words with pictures (Sadoski, 2005), or using imagery methods (e.g., Atkinson, 1975; Cohen & Johnson, 2011). Another technique is the use of gestures, which can be implemented either by having learners observe gestures that are congruent with the meaning of the word that has to be learned, or by having them observe and imitate such gestures.

Gesture observation indeed seems to improve language learning. When parents are trained to use verbal information in combination with iconic gestures (i.e., images of concrete entities and/or actions, for example for the word ‘to mash’ a gesture could be: moving one hand down with force and the other hand open as if holding a bowl; Mcneill & Levy, 1982) during the interaction with their 24 month-olds, these children gain larger vocabularies than infants who did not receive this input (Goodwyn, Acredolo, & Brown, 2000). In a similar vein, when children are requested to learn a number of concrete nouns and descriptive adjectives from their first language, their performance scores are higher when the words were learned in combination with the observation of a gesture that was congruent with the meaning of the word than when words were learned without gesture observation (Tellier, 2005; see also: Valenzeno, Alibali, & Klatzky, 2003). Also, when learning novel locomotion verbs (e.g., to stride; De Nooijer et al., 2014) or a novel proposition (e.g., under; McGregor, Rohlfing, Bean, & Marschner, 2008) gesture observation during learning increases performance scores.

One explanation for the finding that observing a gesture that is congruent with what has to be learned can support word learning comes from research within the field of neuroscience that has shown that observing actions executed by others activates parts of the motor system that would also be activated when the learners would execute this action themselves. Some suggest that this is due to activation of the ‘mirror neuron system’ (Rizzolatti & Craighero, 2004), and although this system is highly debated (e.g., Glenberg, 2011; Gallese, Gernsbacher, Heyes, Hickock, & Iacoboni, 2011), such findings do show that sensorimotor activation (which embodied accounts of cognition would predict to aid memory and learning, see prior section) can be accomplished via gesture observation.

Gesture imitation during word learning

Although gesture observation can improve word learning under certain circumstances, the effects of imitating gestures might be larger than merely observing them, according to several theories. Firstly, action execution has been found to enhance memory traces (e.g., Dijkstra & Kaschak, 2006; Engelkamp & Cohen, 1991). It is assumed that the execution of an action creates a motor trace in memory, which can boost learning performance (Engelkamp & Zimmer, 1984). These findings apply to memory, that is, retaining known information. Learning, which goes beyond memorizing, as it refers to committing novel information to memory, which imposes higher working memory load because information elements have to be organized and integrated, rather than simply rehearsed, has been studied much less.

However, because memory is a prerequisite for learning, studies on the effects of imitation on memory may be predictive for the effect gesture imitation would have on learning.

Although both action observation and imitation activate the motor system, action observation activates the motor system to a lesser degree than action execution (Keysers & Gazzola, 2009). When it is not sensorimotor activation per se, but the strength of the activation that matters, then imitating a gesture would be expected to lead to better learning. Indeed, it appears that word learning improves when children imitate a gesture compared to when they only observe a gesture, when learning words from a first (Tellier, 2005) or second language (Tellier, 2008). Brain research on recognition of words that were previously learned supports the idea that it is especially helpful for word learning when actions are actively performed (James & Swain, 2011; see also: Macedonia, Müller, & Friederici, 2011). fMRI research with children as participants suggests that motor areas of the brain were activated to a larger extent when hearing verbs that were previously learned by actively performing congruent actions, than when the actions were only observed (James & Swain, 2011). These findings indicate that sensorimotor information is reactivated when recognizing words that were previously learned, but only when having actively performed congruent actions, providing evidence for the idea that sensorimotor information may lead to a richer representation of what has to be learned, which could help to improve word learning (Wellsby & Pexman, 2014). Other theories state that it is not so much the motor component in gesture imitation that leads to learning effects, but that the more enriched representation originating from multi-modal input leads to the beneficial effects and deeper representations (Shams & Seitz, 2008).

Whether it is because of the motor component or because the deeper representation that gesture imitation creates, learning novel words via gesture imitation seems to improve learning more than via gesture observation. For example, novel concrete verbs and adjectives were learned better when congruent gestures were imitated compared to when gestures were merely observed. This seems to be true for learning words in a first (Tellier, 2005) or a second language (Tellier, 2008). Moreover, a recent study suggests that the definitions of novel object-manipulation verbs (e.g., to mash) were better learned by nine to eleven year-olds when they were accompanied by gesture imitation compared to gesture observation (De Nooijer et al., 2013).

The effect of gestures on word learning in adults

Although many studies on the effect of gestures in word learning were conducted with children (e.g., De Nooijer et al., 2014; McGregor et al., 2008; Tellier, 2005, 2008), word learning in adults (e.g., in the context of second language learning) might also be improved by gestures. For example, adults learned the meaning of novel Japanese verbs better when a gesture congruent with the meaning of the word was observed during learning (Kelly, McDevitt, & Esch, 2009). Gesture imitation also seems to improve word learning in adults when concrete or abstract words from an artificial language were learned while imitating iconic gestures (Macedonia et al., 2011; Macedonia & Knösche, 2011). Also, when learning new expressions in a second language, learning performance improved when adults imitated congruent gestures, compared to when expressions were learned without seeing the gesture (Allen, 1995).

Even though the use of gestures seems to improve word learning in adults, they might be more beneficial for children than for adults. When children learn new action words they do not know the word, or the definition of that word yet, and they likely also lack a perceptual and motoric mental representation of the meaning. This might make the motoric information conveyed through gestures (particularly though gesture imitation, where the motor system is activated to a larger extent than for gesture observation) especially relevant for children who learn novel words in their first language. When children or adults learn new words in a second language, or when adults learn new synonyms in the first language, they already know the meaning of that word or the action in their first language. They now only need to link the new word to a word or action they have already stored in their mental lexicon. So, when the definition of action words that are already known to the learner needs to be learned, the expectedly larger effect of gesture imitation over observation during learning might disappear, because the gestures may become redundant (and it is known that presenting learners with redundant information does not help, and may even hinder their learning; Kalyuga & Sweller, 2014). When memorizing the definition of a familiar word, gesture imitation might, thus, have little effect compared to gesture observation on learning for both children and adults, because they all have experience with the action. This idea is in line with previous research that suggested that sensorimotor experience might only help learning when this experience is relevant for what has to be learned (Kiefer & Trumpp, 2012). On the other hand, children may still benefit somewhat (though less strongly than when learning new words) from gesture imitation over observation while adults do not. First, adults have more years of perceptual and motor experience with the words (/actions) than children and consequently have more automated representations that come to mind with little effort and require little if any working memory capacity and might, therefore, not have to execute the gestures themselves in order to increase their word learning performance. Second, adults have larger working memory capacity than children, as a consequence of which they may be less dependent on multimodal information presentation (which can increase the functional capacity of working memory by using multiple working memory channels; e.g., Low & Sweller, 2014) than children.

In contrast, when learning a novel action word, where both the word and the action are unknown to the learner, the task is more difficult than when the word and action are already known and only the definition needs to be learned. Previous research has shown that gestures are mostly beneficial in difficult tasks (McNeil, Alibali, & Evans, 2000). It might, therefore, be expected, that the effects of gesture imitation over observation are larger when novel words have to be learned compared to when the definition of known words is learned. When the task is more difficult the richer mental representation created via gestures might lighten the working memory load. Moreover, also when learning the definition of novel action words, differences between children and adults may be expected, as in children the working memory capacity is still developing (Olesen, Macoveanu, Tegnér, & Klingberg, 2007). For this reason, the same task might be more difficult for children than for adults, meaning that the gestures could be expected to be more beneficial to learning the definition of a known or novel word, in children than in adults. As working memory capacity is related to verbal ability (Daneman & Green, 1986), verbal ability might also play a role in the effectiveness of gestures. To illustrate, observing a gesture in combination with verbal input leads to better word learning performance in young children with low language abilities, than for children with high language abilities (Rowe, Silverman, & Mullan, 2013). Also, in adults this seems to be the case, as cueing adults with self-generated gestures during the post-test increased performance of adults with *low* verbal abilities (Frick-Horbury, 2002) On the other hand, there are some indications that a certain amount of working memory capacity (and level of verbal ability) is necessary for gesture imitation to boost learning performance more than gesture observation, as in a recent study only children with high verbal abilities seemed to benefit from gesture imitation compared to observation when learning novel words in the first language (De Nooijer et al., 2014).

The present study

In this study we investigated the effects of gesture observation versus imitation in both children and adults on learning novel or known action verbs. Although the effects of gestures on word learning have been investigated with a number of word classes (e.g., nouns, adjectives and prepositions; e.g., De Nooijer et al., 2013; McGregor et al., 2008; Tellier, 2005), verbs are an especially interesting class of words, given that action verbs have a direct link to the motor system. Prior research using different types of action verbs has suggested that gesture observation and imitation was mainly effective for learning object-manipulation words, in which there is a strong link with the motor system (De Nooijer et al., 2013; De Nooijer et al., 2014), which is why this category of words is under investigation here. The verbs used in this study were either known to the participants or were completely novel. This means that in the first category a verbal definition has to be remembered of a word which is already available in the mental lexicon and of which the action is known, while for the second type a completely new action word and definition have to be learned, with which the learner has no (or very little) motoric experience.

We hypothesized that more verbal definitions would be learned when gesture imitation was used during learning, compared to gesture observation, based on previous literature (e.g., De Nooijer et al., 2013; Tellier, 2005). We expect this to be true for both children and adults. In addition, we expect an interaction between gesture type (observation versus imitation) and word type (known versus novel action verb), in which we hypothesize that gesture imitation will be most effective for learning the definition of a novel word. For these words, the gestures provide additional information, making the gesture information more relevant, which could aid learning (e.g., Kiefer & Trumpp 2012). Moreover, learning a novel word is a more difficult task, and as previous studies have suggested that gestures support learning especially in difficult tasks, we also expect that to be the case here (e.g., McNeil et al., 2000). In addition, the effects of gesture imitation could be stronger for children than for adults, as adults have a larger working memory capacity than children. Given that adults have more perceptual and motor experiences with a large array of actions, they might still be able to connect the novel action to known actions that resemble the new one, and activate these representations. As a consequence they might not need the extra motoric information gesture imitation provides them with. Lastly, although theoretically less interesting, we expect adults to score better than children and performance on known verbs to be higher than on novel words.

Method

Participants

Children. Participants were 59 Dutch primary school children (22 male, 37 female), between nine and eleven years of age (*M* = 10.1, *SD* = 0.6), who were recruited from three primary schools in the same area of the Netherlands. We tested primary school children of approximately ten years old because these children are still in the process of acquiring new words in their first language, while being old enough to be able to participate in a study that requires sustaining attention for approximately 20 minutes, understanding the instructions, and learning a substantial number of new words within a short period of time. Children were awarded a small gift for participation and were pseudo-randomly assigned to the gesture observation (*n* = 29) and gesture imitation conditions (*n* = 30), matched for language ability according to a standardized vocabulary test.

Adults. Participants were 60 university students (15 male, 45 female) with an average age of 21 years (*SD* = 2.83). They participated in exchange for course credit or a small monetary reward. The experiment lasted approximately 15 minutes. Participants were randomly assigned to the gesture observation (*n* = 29) or gesture imitation (*n* = 31) conditions.

Materials

Verbs. Material consisted of 12 action verbs, six of which were known to both the children and the adults and six were novel to both groups. The known words were actions that the children had come into contact with and probably had motor experience with (e.g., molding, typing or tying). To test whether the novel words were indeed novel to adults, we pretested 15 of these verbs. A total of 19 fourth-year psychology students (14 female, 5 male) indicated whether they knew any of these 15 words. On the basis of these pretest results, six words were selected for the actual experiment. None of these words were known to any of the participants in the pretest. In addition, these were all words that participants were unlikely to have had motoric experience with (e.g., the Dutch word "hieuwen", meaning to hoist or haul in the anchor of a ship). The known and novel verbs were matched for phoneme length (6.33 vs. 6.66 respectively). Lexical frequency was determined by the mean logarithmic lemma frequency per million for each verb type using the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). The average frequency for the known words was 1.059 (*SD* = .55) versus 0 for the novel words.

Gestures. For each verb a congruent gesture was created. This gesture was made by a young adult female actor. Our choice for using an adult model was motivated by considerations of ecological validity and practical implications, meaning that in the classroom, gestures during instructions will typically be made by a teacher (and in Dutch primary education, the majority are female teachers). All actions were depicted as realistically as possible and might therefore be described as pantomimes. All gestures were manual, executed with both hands. The videos lasted for six seconds.

Verbal definitions. The verbal definition of the word was always presented twice; the video of the gesture was shown only during the second presentation. The verbal definitions were audio recorded by the same female actor. In the gesture videos this audio file was overlaid on the gesture video.

Posttest. The posttest consisted of a recall test, in which the children and adults were asked to provide a definition of the verbs. This was a paper-and-pencil test, but to the children, the target words were read orally and they answered orally. They were then asked to write down what they answered orally. On this written answer, they were eventually scored. This method was chosen to prevent possible reading difficulties from interfering with children's test performance. Adults read and filled out the posttest themselves.

Procedure

Before the experiment, scores on a standardized vocabulary test were collected from the schools for all children participating in the study. Scores on this vocabulary test consist of five categories, which have cut-off scores based on a national norm group: A (i.e., score in the top 25% of scores nationally), B (i.e., score in the next 25% range, just above the national average), C (i.e., score in the next 25%, just below the national average), D (i.e., score in the next 15%, scoring well below the national average), E (i.e., score in the 10% lowest scores). Children were considered to have high verbal abilities when they scored an A or B (N = 13; 5 in the observation condition, 8 in the imitation condition) and low verbal abilities when they scored a C, D or E on this test (N= 39, 22 in the observation condition, 17 in the imitation condition). Of 7 children the verbal abilities were unknown as they just transferred to a different class and therefore their scores on the standardized test were not known to the teacher.

In individual sessions of approximately 20 min. duration, the children and adults learned the definitions of 12 verbs in two blocks. Participants were told that they would be learning the definition of a number of words, some of which they might already know. In each case the task was to remember the definition of the word as they heard it in the audio file. The task looked exactly the same in the observation and imitation conditions, only the instructions differed. Participants were first presented with the written word for two seconds. For the children this word was read aloud by the experimenter while the adults read the words they had to learn themselves. Participants were then presented with the verbal definition of the word twice. During the second presentation participants also saw the gesture congruent with the meaning of the word. Both definition and gesture video lasted for six seconds.

Participants were only instructed to watch and listen to the meaning of the words and to try and remember them in the observation condition. In the imitation condition, participants were instructed to immediately imitate the gesture they saw. Words were presented in two blocks of six words containing both novel and known words. Words were randomized within blocks and the blocks themselves were counterbalanced, resulting in the use of two lists for the observation condition and two lists for the imitation condition. The children completed a posttest over the first six words and a second posttest over the last six words, while the adults completed one posttest over the 12 words. The task might otherwise have been too simple for the adults, which could result in ceiling effects.

Data Analysis

The recall test was scored by awarding one point for each correct answer. In the recall task participants could also be rewarded a half point when one element of the definition was missing (e.g., ‘to squeeze something’ instead of ‘to squeeze something to give it shape’ for ‘to mold’). Ten percent of the data was coded by two coders. The second coder was blind to the experimental condition of each response. The interrater reliability was Kappa = .90. Rater agreement can, therefore, be considered high.

Results

Mean scores and standard deviations for the recall test are given in Table 1. Data were analyzed with a Repeated-Measures ANOVA with word type (known, novel) as within-subjects factor and gesture condition (observation vs. imitation) and age (children vs. adults) as between-subject factors. There was a main effect of word type, $F(1, 115) = 279.69$, $p < .001$, $\eta_p^2 = .71$, indicating that, as expected, overall performance on the known verbs was significantly higher than on the novel verbs (see Table 1). There was also a main effect of age, $F(1, 115) = 82.65$, $p < .001$, $\eta_p^2 = .42$, with, as expected, the adults scoring higher than the children on the recall task (see Table 1). In contrast to our hypothesis, there was no main effect of gesture condition, $F(1, 115) < 1$, $p = .621$, $\eta_p^2 = .002$, nor a significant gesture condition x word type interaction, $F(1, 115) < 1$, $p = 8.48$, $\eta_p^2 < .001$. However, there was a small but significant interaction between age and gesture condition, *F*(1, 115) = 4.36, *p* = .039, η_p^2 = .04. No other interactions were significant (i.e., word type x age, *F*(1, 115) < 1, *p* = .934, η_p^2 < .001; word type x age x gesture condition, *F*(1, 115) < 1, *p* = .724, η_p^2 = .001). As a large number (39 out of 59) of children had low verbal abilities, it was not possible to investigate whether verbal ability plays a role in the effectiveness of gesture imitation over observation in children. [1]

Table 1. Means (and Standard Deviations) of % Items Correct on the Posttest

	Condition			
	Observation		**Imitation**	
Verb type	*Children (n = 29)*	*Adults (n =29)*	*Children (n = 30)*	*Adults (n = 31)*
Known verb	56.89 (19.25)	88.22 (11.68)	60.27 (24.83)	81.18 (17.48)
Novel verb	20.40 (16.90)	52.87 (20.93)	26.11 (20.26)	45.16 (21.17)
Overall score	**38.79 (15.31)**	**70.55 (13.22)**	**43.19 (17.46)**	**63.17 (16.07)**

To determine what the interaction between age and gesture condition signified, we first visually inspected the data, which suggested that the interaction between age and gesture type indicates that there is a benefit of gesture imitation over gesture observation for children, while it is the other way around for adults. To test this we first performed an ANOVA for the data of the children with gesture condition as independent variable and performance scores (on both novel and known words) as dependent variable. The ANOVA, however, did not show a significant effect of gesture condition for the children, *F*(1, 57) = 1.14, *p* = .289, η_p^2 = .02.Similarly, for the adults there was no significant effect of gesture condition, *F*(1, 58) = 3.74, *p* = .058, η_p^2 = .06.

[1] When analyzing the data for the children with low verbal abilities there was no main effect of gesture condition, *F*(1, 37) = 3.06, *p* = .089, η_p^2 = .076 (where numerically children score better under the gesture imitation condition), nor a significant interaction between gesture condition and word type, *F*(1, 37) = 1.29, *p* = .264, η_p^2 = .034. These findings are in line with findings by De Nooijer et al. (2014), where children with low verbal abilities also did not seem to benefit from gesture imitation during word learning.

Discussion

Research has shown that gestures can support word learning (e.g., De Nooijer et al., 2013; Tellier, 2005). However, it is still unclear which factors play a role in the effectiveness of gestures; the level of motor activation, the familiarity of the word of which the definition has to be learned, or the age of the learner (because of age-related differences in working memory). In this study we investigated one factor that might influence the effectiveness of gesture use during word learning, namely the novelty of the word of which the definition has to be learned. We investigated the effects of gesture observation versus gesture imitation in both children and adults on learning the definition of known or novel words. Although word learning performance has been found to increase in both children and adults when gestures are used during learning (e.g., De Nooijer et al., 2013, Kelly et al., 2009, Macedonia & Knösche, 2011; Tellier, 2005), little is known about any differences between the use of gestures in children's and adult's word learning. To the best of our knowledge this is the first study that has directly compared the effects of gestures during word learning in children and adults. This study could provide valuable insight as to when gestures can benefit learning.

As expected, we found a main effect of age, in which adults scored higher on the word-learning task than children. In addition, there was a main effect of word type, showing that the definitions of known action verbs were remembered better than those of novel action verbs. These findings are, however, quite obvious and not that interesting theoretically, so they will not be further discussed. Surprisingly, and in contrast with our hypotheses, there was no main effect of gesture condition. Based on the literature we expected a benefit of gesture imitation over gesture observation for both children and adults (e.g., De Nooijer et al., 2013; Macedonia et al., 2011; Macedonia & Knösche, 2011, Tellier, 2005; 2008). It seemed that for children, there was no difference between gesture imitation and gesture observation, although they numerically scored somewhat higher under the imitation condition on both the known and the novel words. For the adults, in contrast, gesture observation seemed to lead to higher performance scores than gesture imitation; although this direct comparison also failed to reach statistical significance (p = .058).

So how can we explain these findings? Let us first look at the performance of the children. In a previous study with a similar design, similar verbs and participants of a similar age (De Nooijer et al., 2013) a beneficial effect of gesture imitation over observation was found. Although, as mentioned before, children in this study seemed to do better under the gesture imitation condition, which would be in line with De Nooijer et al., 2013. The effects were, however, very small, as the interaction between gesture type and age was significant, but follow-up post-hoc tests were not. One reason for why the effects in this study were smaller, might be that the participants consisted of relatively more children with low verbal skills (high verbal ability: *N* = 13, low verbal ability: *N* = 39) as indicated by the same standardized vocabulary test) compared to the children that participated in the study by De Nooijer et al., 2014. (high verbal ability: *N* = 21, low verbal ability: *N* = 23) or in the study by De Nooijer et al., 2013 (high verbal ability: *N* = 51, low verbal ability, *N* = 64).Verbal ability might play a role in the effectiveness of gestures in word learning (e.g., Frick-Horbury, 2002). A study with children that were one to two years younger found that imitation of object-manipulation words only improved word learning in children with high verbal skills (De Nooijer et al., 2014; see also: Post, Van Gog, Paas, & Zwaan, 2013). This could be one explanation for why a study with a similar design found stronger effects of gesture imitation, than we did. Moreover, an explorative analysis on the current data showed that when only using the data of the children with low verbal abilities, there was no beneficial effect of using gesture imitation over observation. As there was only a very limited number of children with high verbal abilities we could not test, whether in line with De Nooijer et al., 2014, children with high verbal abilities do seem to benefit from gesture imitation during word learning. In contrast to what might be expected on the basis of literature that found that gestures mostly benefit learners with low verbal ability (e.g., Rowe et al., 2013; Frick-Horbury, 2002), this study is not in line with those findings. Although working memory load might be reduced, when using gestures, it could be the case that the added difficulty of simultaneously learning a word and thinking about how to imitate it might, however, have been too challenging for these children with low verbal skills and could have caused the poorer performance in the action imitation condition (de Nooijer et al., 2014). If our children had lower working memory capacity, this could mean they had more trouble executing the gestures while listening to the meaning of the words, therefore experiencing less benefit from the gestures, as vocabulary knowledge has been found to relate to working memory capacity (Daneman & Green, 1986). This means that there is a crucial time period in which children can start to use gesture imitation to their benefit in learning new words and most likely this means that their verbal ability has to have reached a certain level.

In this study we were also interested in whether adults can benefit from gestures in a similar way to children, when they learn known and novel words in their first language. The adults, however, did not seem to benefit from gesture imitation, and scored substantially, though not significantly, better under the gesture observation condition than under the gesture imitation condition. What makes this study different from other studies in the field that did find a beneficial effect of gesture imitation on learning words in adults? Only a few studies investigated this issue in adults and most used second language learning (e.g., Allen, 1995; Kelly et al., 2009 Macedonia et al., 2011; Macedonia & Knösche, 2011). Other studies on gesture use and language in adults in the *first* language, have mostly looked at language comprehension and not word learning (e.g., Kelly, Barr, Church, & Lynch, 1999). Kelly et al. (1999), for example, showed that adults understood an utterance better when it was combined with an iconic gesture. They understood an utterance to be an indirect request more often when the utterance was combined with a relevant pointing gesture. However, this reflects very different processes than learning new words in the first language, or memorizing utterances. Studies on the enactment effect did find memory advantages when participants acted out the action from the sentence (e.g., Engelkamp & Zimmer, 1997), however here adults had to create their own action and not imitate something they saw. This might make the task more active and could increase the effect of action execution on learning. Self-involvement could, therefore, be a crucial factor for the potential effects of using gestures during learning in adults (e.g., Kormi-Nouri, 1995).

Another reason for why gesture imitation might not improve word learning in the first language could be that the task was too easy, as gestures mostly benefit learning in difficult tasks (McNeil et al., 2000). Scores on especially the novel words were, however, not that high and therefore this explanation seems unlikely. So, this study suggests that using gestures during word learning is less effective in adults than in children.

In sum, this study suggests that gesture imitation is not necessarily more effective than observation, and that for adults, observation might even be more effective. The effectiveness of gestures during word learning should for this reason be assessed in the light of the stage of development of the learner.

▪ Chapter 5

When left is not right: Handedness effects on learning object-manipulation words using pictures with left- or right-handed first-person perspectives

This chapter has been published as:

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives. *Psychological Science, 24,* 2515-2521. doi: 10.1016/j.actpsy.2013.05.013.

Abstract

According to the body-specificity hypothesis, hearing action words creates body-specific mental simulations of the actions. Handedness should, therefore, affect mental simulations. Given that pictures of actions also evoke mental simulations and often accompany words to be learned, would pictures that mismatch the mental simulation of words negatively affect learning? We investigated effects of pictures with a left-handed, right-handed, or bimanual perspective on left- and right-handers' learning of object-manipulation words in an artificial language. Right-handers recalled fewer definitions of words learned with a corresponding left-handed-perspective picture than with a right-handed-perspective picture. For left-handers, there was no effect of perspective. These findings suggest that mismatches between pictures and mental simulations evoked by hearing action words can negatively affect right-handers' learning. Left-handers, who encounter the right-handed perspective frequently, could presumably overcome the lack of motor experience with visual experience and, therefore, not be influenced by picture perspective.

Introduction

According to theories of embodied cognition, language, perception, and action are closely intertwined (e.g., Barsalou, 1999; Glenberg, 1997; Stanfield & Zwaan, 2001). Studies have shown, for instance, that when hearing a sentence such as “Close the drawer,” the denoted action interacts with the subsequent performance of an action (e.g., Glenberg & Kaschak, 2002). This interaction seems to be based on internalized personal experience with the denoted action. For example, people remember high-manipulability words (e.g., *camera*) better than low-manipulability words (e.g., *table*) because high-manipulability words evoke greater activation in motor regions of the brain (Madan & Singhal, 2012). Moreover, research has shown that when imagining manual-action verbs, such as *writing*, right-handers activate the left premotor cortex, whereas left-handers activate right premotor areas (Willems, Hagoort, & Casasanto, 2010). In addition, right-handers react faster to manual-action verbs on a lexical decision task after stimulation of the left premotor cortex than after stimulation of the right premotor cortex, which suggests a causal link between activity in the premotor cortex and understanding of action verbs (Willems, Labruna, D’Esposito, Ivry, & Casasanto, 2011). These findings suggest that left- and right-handers may process object-manipulation words differently because they have had different interactions with these actions (e.g., Willems et al., 2011). This idea is in line with the body-specificity hypothesis (Casasanto, 2009), which states that reading or hearing about an action leads to the creation of body-specific mental simulations of the denoted action.

Research has also shown that seeing an object or a picture of an object automatically activates motor information (Borghi et al., 2007; Sumner et al., 2007), but this effect might differ for left- and right-handers. For example, right-handers are sooner inclined to purchase a displayed product when it is presented in an orientation that is suitable for use by a right-handed person. Orienting a product toward the dominant hand, such as by showing a picture of a mug with the handle on the right, facilitates simulation of using the product for right-handers, which affects behavioral intentions (Elder & Krishna, 2012). These results suggest that right-handers simulate right-handed actions, whereas left-handers might create left-handed simulations. On the basis of the body-specificity hypothesis, we predicted that left-and right-handers should make different mental simulations and should therefore be differentially influenced by seeing an action performed by a left hand or a right hand.

In sum, ample evidence has suggested that when a person hears actions described or sees depictions of actions or objects, body-specific mental simulations are made. An interesting question is whether mismatches between the mental simulation evoked by a verbal action description and the mental simulation evoked by a picture that depicts the action would hamper learning and, if so, whether this effect would be the same for left- and right-handers. Because left-handers are frequently required to use their right hand (e.g., Stins, Kadar, & Costall, 2001), they might be less affected by such a mismatch. For instance, a recent study by Apel, Cangelosi, Ellis, Goslin, and Fischer (2012) showed that motor activation induced by seeing pictures influenced memory performance. In the study, right-handers remembered more instructions when the picture presented an object's handle oriented to the right and actions had to be performed with the right hand. Although this result suggests that motor simulation can support sequential memory for actions, this effect was not found for left-handers.

Left-handers live in a right-handed world, given that only approximately 10% of the world population is left-handed (Coren & Porac, 1977). Therefore, left-handers are often forced to use their right hand, whereas right-handers do not have to use their left hand very often (e.g., Stins et al., 2001). Moreover, just like right-handers, left-handers see more people in their environment using their right hand than using their left hand. Left-handers might therefore not have strong hand-specific embodied connections. Indeed, left-handers show mirror neuron–system activity for action execution with both their dominant and their nondominant hand, whereas right-handers show this activation only for their dominant hand (Rocca, Falini, Comi, Scotti, & Filippi, 2008). This finding is also supported by other research (e.g., Willems & Hagoort, 2009) and is consistent with the body-specificity hypothesis, which posits that word meanings are in part made up of simulations of one's own actions. Right-handers would therefore make a right-handed simulation in response to a description of an action, whereas left-handers might not always create left-handed simulations, given that they use their right hand for manual actions as well and have extensive visual experience with right-handed actions.

In our three experiments, we sought to ascertain whether there is an effect of mental simulations in word learning for left- and right-handers by using pictures of actions congruent with the meaning of object-manipulation verbs. Word learning is an important part of second-language acquisition, and new words can successfully be learned using pictures (e.g., Tellier, 2008). However, to our knowledge, the role of mental simulation in word learning with pictures has not been investigated. According to the body-specificity hypothesis, concepts are constituted by mental simulations, which are based on people's own perceptuo-motor experiences. For this reason, seeing a picture from a perspective incongruent with one's own mental simulation might hinder the encoding and consolidation process of the new word.

On the basis of our review of the literature, we predicted that there would be an interaction between (a) the picture perspective shown to a participant when he or she learned the meaning (i.e., definition) of a new (artificial) word denoting a verb and (b) the handedness of the participant. We expected that right-handers' recall would be negatively affected by seeing a left-handed perspective (L) compared with a right-handed perspective (R) during learning (i.e., recall would be greater for words paired with R pictures relative to L pictures: R > L) because this perspective would mismatch their own mental simulation. In contrast, we expected that left-handers either would not be affected by seeing the right-handed perspective (i.e., recall would not differ between words paired with L and R pictures: L = R) or at least would not be affected as much as right-handers when they saw the left-handed perspective (i.e., recall would be greater for words paired with L pictures relative to R pictures—L > R—but the effect would be smaller), given that left-handers live in a right-handed world.

Pictures with a bimanual perspective (B), showing both hands performing an action, should be body specific for both right-handers and left-handers. Therefore, we expected that right-handers would not be affected by seeing the bimanual perspective (i.e., recall would not differ for words paired with B and R pictures: B = R) and that both words paired with B pictures and words paired with R pictures would be recalled better than words paired with L pictures (i.e., B > L and R > L). For left-handers, we expected that seeing the bimanual perspective would have no effect (i.e., recall would not differ for words paired with B and L pictures: B = L) and either that words paired with B or L pictures would be recalled better than words paired with R pictures (B > R and L > R) or, if left-handers were not hampered by the right-handed perspective because of their experience with it, that seeing either the B or the R perspective would have no effect (i.e., there would be no difference in recall across conditions: B = L = R).

To assess the reliability of our results from Experiment 1, we performed a direct replication of this study approximately 3 weeks later with a different sample of participants (Experiment 2). In addition, to control for possible item effects, we conducted a follow-up study with right-handers only, in which the artificial words were rotated across definitions (Experiment 3).

Method

Participants

Participants were recruited via Amazon's Mechanical Turk (see Buhrmester, Kwang, & Gosling, 2011; Zwaan & Pecher, 2012). All participants were U.S. residents and native speakers of English. They received $0.75 in return for their participation in the task, which required approximately 15 to 20 min to complete. The sample for Experiment 1 comprised 60 left-handed (27 female, 33 male) and 60 right-handed (33 female, 27 male) adults (mean age = 33.6 years, *SD* = 11.9). In Experiment 2, 60 left-handed (33 female, 27 male) and 60 right-handed (35 female, 25 male) adults participated (mean age = 29.3 years, *SD* = 9.5). In each of the first two experiments, we also collected data from numerous other right-handed participants. Here, however, we report data from only the first 60 right-handed participants in each experiment to equate the sample sizes for right- and left-handers (but the same pattern of results was obtained analyzing the full sample of right-handers from Experiment 1). In Experiment 3, 120 right-handed adults (72 female, 48 male; mean age = 33.9 years, *SD* = 11.7) participated via Mechanical Turk.

Materials

Materials consisted of 18 words from the artificial language "Vimmi" (Macedonia & Knösche, 2011). We used an artificial language to avoid any influence of participants' possible prior knowledge of the target language and any possible cognates that might exist in the two languages, which might influence word learning. Words were matched for length, and each word was randomly coupled with the definition of a manual object-manipulation verb (e.g., *luko* was coupled with the definition of "to pour," i.e., to dispense from a container). There were 12 definitions of unimanual object-manipulation words, each of which was coupled with a picture that showed a hand executing the denoted action from either an L or an R first-person perspective (source pictures were mirrored horizontally to create otherwise identical L and R perspectives; see the top row of Fig. 1 for example images). We coupled 6 definitions of bimanual object-manipulation words with first-person-perspective pictures in which both hands were used to execute the action, making the perspective similar for left- and right-handers (see the bottom row of Fig. 1 for an example image).

In Experiments 1 and 2, each artificial word was always coupled with the same definition. In Experiment 3, the artificial words were rotated across definitions (i.e., for different participants, *luko* would have different definitions) to control for possible item effects.

Design and procedure

The recall test asked participants to provide the English definition of the Vimmi words as literally as possible.[1] In a within-subjects design, participants learned each of the 18 Vimmi words by listening to the definition and viewing a picture in three conditions: 6 unimanual words presented with an R picture, 6 unimanual words presented with an L picture, and 6 bimanual words presented with a B picture. Participants learned 6 words in each of three

[1] We also tested participant's ability to provide an English translation for each artificial word; however, we had no a priori hypotheses about the effects of pictures' handedness perspective on translation, and the results were inconsistent across studies and show no such effects. Therefore, we have not reported these data here, but they can be obtained from the corresponding author.

blocks in which two pictures from each condition were presented, and were tested after each block.

Type of action	Word	Definition & Picture
Unimanual	Luko (“pour”)	To dispense from a container  or 
Bimanual	Lefa (”type”)	To produce text using a keyboard 

Figure 1. Example materials. The top row shows a unimanual object-manipulation word coupled with its definition and left- and right-handed first-person perspective pictures of the denoted action, whereas the bottom row depicts a bimanual object-manipulation word coupled with its definition and a first-person perspective pictures of the denoted action.

In each block, participants heard a definition of each word they had to learn, during which the Vimmi word was also presented on the screen. After the definition was presented twice, the program automatically proceeded to the next word. During the second presentation of the definition of the word, a picture was shown in which the denoted action was being executed.

In Experiments 1 and 2, to avoid order effects, we counterbalanced words across the three blocks such that all words appeared in all blocks. L and R pictures were also counterbalanced, resulting in six lists. In Experiment 3, L and R pictures were still counterbalanced, but all Vimmi words rotated in the six lists across the three blocks.

We collected data on participants' handedness, demographics, and thoughts about the study at the end of the experiments because we did not want to provide cues about the goal of the study. None of the participants correctly guessed our hypothesis. Because we determined participants' handedness only after the experiment and because left-handers compose only approximately 10% of the world population (Coren & Porac, 1977), the total sample was 593 in Experiment 1 (60 left-handers and 533 right-handers) and 613 in Experiment 2 (60 left-handers and 553 right-handers). Data from the first 60 right-handers in both experiments were scored and analyzed in comparison with the respective left-handers' data. Nonnative speakers of English and participants who indicated that they were in a highly noisy and distracting environment while participating (i.e., people who scored a 7 or higher on a 9-point scale for background noise and distraction) were removed and replaced by new participants prior to further analyses. This procedure resulted in the exclusion and replacement of 5% of participants in Experiment 1 and 7.5% of participants in Experiment 2. In Experiment 3, which included only right-handers, 19% of participants were replaced (because all left-handers had to be excluded as well).

Scoring

Recall performance was scored by awarding one point for each correct answer. Participants could be rewarded a half point if one element of the definition was missing (e.g., if a participant responded with "make clothing" instead of "to make clothing using needles" for "'to knit"). We had 10% of the data scored by a second rater who was blind to the experimental conditions. Interrater reliability was high (κ = .82).

Results

Figure 2 displays participants' recall results from all three experiments. Recall scores from Experiments 1 and 2 were analyzed with 3 × 2 mixed-model analyses of variance with condition (L, R, B) as the within-subjects variable and handedness of the participant as the between-subjects variable. In Experiment 1, the interaction between condition and handedness was significant, $F(2, 118) = 3.54$, $p = .03$, $\eta_p^2 = .03$. Paired *t* tests with Bonferroni-adjusted alpha levels of .017 showed that right-handed individuals displayed the expected pattern—R > L: $t(59) = 2.58$, $p = .01$, $d = 0.25$, 95% confidence interval (CI) = [–0.12, 0.60]; B > L: $t(59) = 4.06$, $p < .001$, $d = 0.39$, 95% CI = [0.02, 0.74]; B = R: $t(59) = 1.57$, $p = .12$, $d = 0.14$, 95% CI = [–0.22, 0.50]. For the left-handers, results showed a contrasting pattern—L = R: $t(59) < 1$, $p = .49$, $d = 0.07$, 95% CI = [–0.29, 0.43]; B = L: $t(59) < 1$, $p = .38$, $d = 0.08$, 95% CI = [–0.28, 0.44]; B = R: $t(59) = 1.42$, $p = .16$, $d = 0.15$, 95% CI = [–0.21, 0.51].

In Experiment 2, the interaction was replicated, $F(2, 118) = 3.56$, $p = .03$, $\eta_p^2 = .03$. Again, paired *t* tests with Bonferroni adjustment showed the predicted pattern for right-handers—R > L: $t(59) = 2.85$, $p = .006$, $d = 0.29$, 95% CI = [–0.08, 0.64]; B > L: $t(59) = 3.02$, $p = .004$, $d = 0.28$, 95% CI = [–0.08, 0.64]; B = R: $t(59) < 1$, $p = .96$, $d = 0.01$, 95% CI = [–0.35, 0.36]. For the left-handers, we again found no effect of seeing the right-handed perspective—L = R: $t(59) < 1$, $p = .63$, $d = 0.06$, 95% CI = [–0.29, 0.43]; however, our other results contrasted with those of Experiment 1—B > L: $t(59) = 2.82$, $p = .006$, $d = 0.29$, 95% CI = [–0.08, 0.64]; B > R: $t(59) = 3.26$, $p = .002$, $d = 0.36$, 95% CI = [0.003, 0.72].[2] In Experiment 3, paired *t* tests with Bonferroni adjustment again showed the predicted pattern—R > L: $t(119) = 2.45$, $p = .01$, $d = 0.18$, 95% CI = [–0.07, 0.43]; B > L: $t(119) = 2.83$, $p = .005$, $d = 0.25$, 95% CI = [–0.01, 0.50]; B = R: $t(119) = 0.99$, $p = .32$, $d = 0.07$, 95% CI = [–0.18, 0.33].

[2] We also scored and analyzed all right-hander's data from Experiments 1 and 2. In Experiment 1, there were 533 right-handers, of whom 473 met inclusion criteria We also scored and analyzed all right-handers' data from Experiments 1 and 2. In Experiment 1, there were 533 right-handers, of whom 473 met inclusion criteria; the patterns of results did not differ when these participants were included in our analyses—R > L: $t(472) = 3.37$, $p = .001$, $d = 0.12$, 95% CI = [–0.01, 0.25]; B > L: $t(472) = 4.74$, $p < .001$, $d = 0.17$, 95% CI = [0.04, 0.29]; B = R: $t(472) = 1.27$, $p = .21$, $d = 0.05$, 95% CI = [–0.08, 0.17]. In Experiment 2, there were 553 right-handers, of whom 511 met inclusion criteria; the patterns of results did not differ when these participants were included in our analyses—R > L: $t(510) = 2.91$, $p = .004$, $d = 0.11$, 95% CI = [–0.01, 0.23]; B > L: $t(510) = 4.59$, $p < .001$, $d = 0.17$, 95% CI = [0.05, 0.29]; B = R, $t(510) = 1.78$, $p = .07$, $d = 0.06$, 95% CI = [–0.06, 0.19].

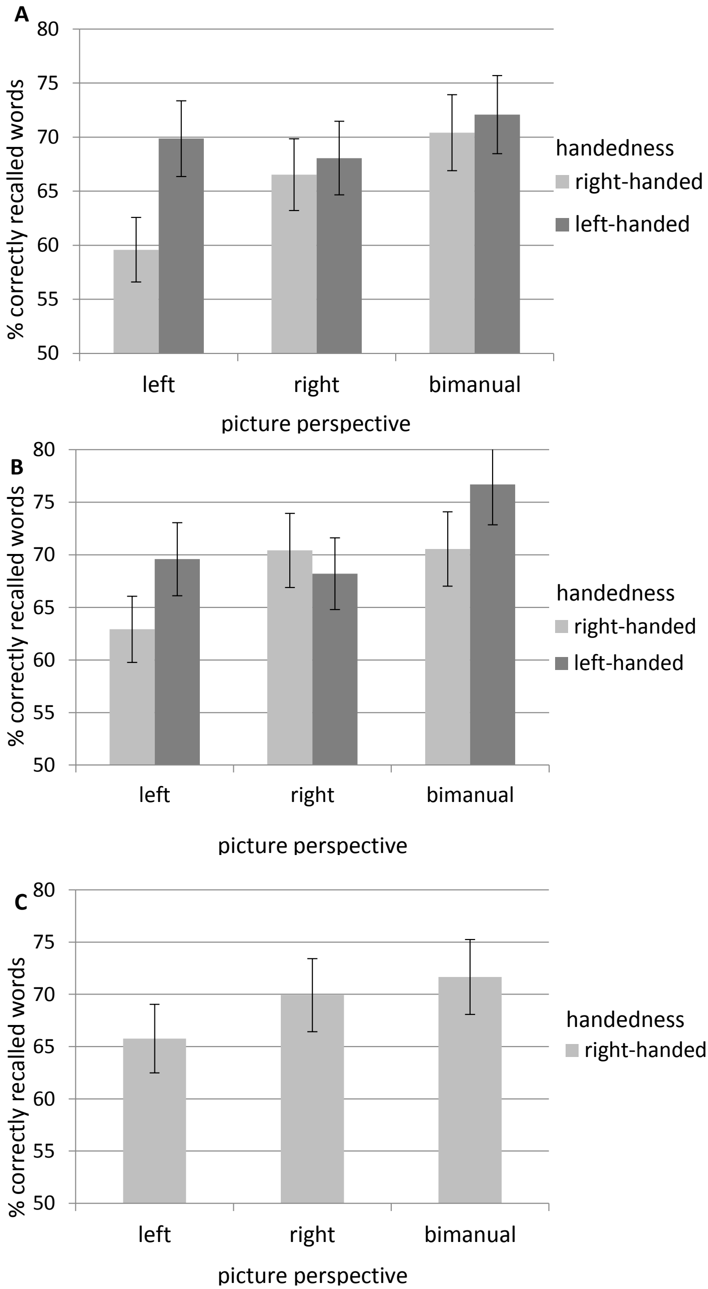


Figure 2. Mean recall performance (percentage of words correctly recalled) as a function of picture perspective and participant's handedness. Results are shown separately for (a) Exp. 1, (b) Exp. 2, and (c) Exp. 3. Error bars represent 95% within subjects confidence intervals calculated using the formula developed by Masson and Loftus (2003).

Discussion

We investigated whether there was an effect of mental simulations in word learning for left- and right-handers by teaching people action verbs in an artificial language using pictures that would match or might mismatch their mental simulation. Our hypothesis that right-handers' learning would be negatively affected by a mismatch between their (right-handed) mental simulation evoked by hearing a verbal definition of an action word and the left-handed picture perspective they saw was confirmed across the three experiments.

We did not find this mismatch effect for the left-handed participants. This differential finding for right- and left-handers is consistent with the results of a study by Apel et al. (2012), which showed that right-handers but not left-handers correctly recalled and executed more instructions for moving objects when their handles were on the same side as their dominant hand. Similarly, right-handers but not left-handers perceived tools that were difficult to grasp as being farther away than easy-to-grasp tools (Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009).

Left-handers live in a right-handed world and, therefore, have visual experience with a right-handed perspective and are also often confronted with objects meant for right-handed use, which forces them to resort to using their right hand (although this is not true for all actions; e.g., writing). Left-handers see the right-handed first-person perspective more often than right-handers see the left-handed first-person perspective. For this reason, left-handers might not have very strong hand-specific embodied representations. Right-handers presumably do have strong hand-specific embodied representations, which is why they mentally simulate actions with a right-handed perspective. When right-handers are shown a left-handed picture perspective, it mismatches their own mental simulation. This effect is consistent with the body-specificity hypothesis, which states that understanding action words requires mentally simulating one's own actions. When a picture mismatches one's own mental simulation, this mismatch might hinder the process of encoding the new information, which could result in lower recall scores. These findings also provide evidence for the embodied view of cognition, which states that memory content is associated with sensory and motor representations of an action and that the mental simulation of an action is part of its representation (Barsalou, 2008). Future research could investigate whether left-handers' learning words that denote actions for which they exclusively use their dominant hand results in a detrimental effect of observing a right-handed compared with a left-handed picture perspective.

In conclusion, this study provides new insight into the effect of mental simulation on word learning, that is, learning the meaning of an unfamiliar word by associating it with a known action. Furthermore, we compared left- and right-handers on this matter, which to our knowledge has never been done. We have shown that viewing a picture of an action that mismatches a right-hander's own mental simulation can negatively influence his or her word learning. An interesting question for future research would be whether the same effect could be found for right-handers when definitions of novel words are coupled with a picture of a third-person perspective, given that right-handers have some visual experience observing left-handed people performing such actions. Furthermore, it would be interesting to investigate whether the effects found in the present study are limited to recall or would persist on a comprehension test. With this study, we have taken a first step toward answering these and related questions regarding the influence of mental simulations on memory and learning.

▪ Chapter 6

Effects of pictures with left- or right-handed first-person and third-person perspectives on learning object-manipulation words

This chapter has been submitted for publication as:

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *Effects of pictures with left- or right-handed first-person and third-person perspectives on learning object-manipulation words.* Manuscript in revision.

Abstract

Prior research showed that for right-handers, left-handed first-person perspective pictures negatively affect word learning compared to right-handed first-person perspective pictures, presumably because the left-handed pictures mismatched the mental simulation of action words. In three experiments we followed up on this finding and investigated whether (1) it would also apply to third-person perspective pictures, (2) we could replicate the effect with first-person perspective pictures, and (3) pictures would improve performance compared to learning new words with verbal definitions alone. Although we replicated the effect with first-person perspective pictures, there was no effect of hand perspective with third-person perspective pictures; presumably, no mental simulation mismatch is experienced with such pictures. Lastly, first- or third-person left- and right-handed perspective pictures did not improve word learning compared to learning without pictures. This study suggests that pictures do not necessarily help in word learning, but when they are used, instructional designers should be aware of handedness effects when using first-person perspective pictures.

Introduction

Word learning is an important part of language acquisition, which is probably why researchers have investigated ways to optimize this process. According to a dual-coding view on vocabulary learning (Sadoski, 2005), verbal information combined with pictorial information (i.e., multimedia learning; Mayer, 2014), can be more effective for learning new words than verbal information alone. However, even though some studies seem to suggest that combining text with pictures enhances performance (e.g., McGregor, Sheng, & Ball, 2007; Smith, Stahl, & Neil, 1987) others do not (e.g., Acha, 2009). The present study investigated the effects of different types of pictures on learning action verbs.

Using pictures in combination with text while learning action verbs might enhance performance, because the pictures provide additional information that might make it easier to consolidate the new information. This is similar to why it is easier to remember action-related sentences when the actions described in the sentence are self-generated during encoding. The additional motor information gained from executing the action enhances memory traces (i.e., the enactment effect; see review by Engelkamp, 1998). Not only action execution, but also implicit action, in the form of pre-motor activation, might affect memory. When learning a new action verb via a verbal definition in combination with seeing a picture of this action, this evokes a mental simulation, according to theories of embodied cognition. Mental simulations are "the reenactment of perceptual, motor, and introspective states acquired during interaction with the word, body, and mind'' (Barsalou, 2008, p.618), which can be evoked both by verbal and visual information (e.g., pictures) and result in pre-motor activity (e.g., Willems, Toni, Hagoort, & Casasanto, 2010). For example, it has been shown that during language comprehension mental simulations are made of what was heard or read. When reading a sentence such as "John put the pencil in the cup", implying a vertical orientation of the pencil, participants are faster to indicate that the word 'pencil' occurred in the sentence when first shown a picture of a pencil in a vertical orientation, which matched the orientation implied by the sentence, than when first seeing a picture of a pencil in a horizontal orientation. This suggests that the participants mentally simulated the object in the implied orientation (Stanfield & Zwaan, 2001; see also: Barsalou, 1999; Glenberg & Kaschak, 2002; Zwaan & Pecher, 2012). Furthermore, seeing (a picture of) an object or tool has also been found to automatically elicit motor information (Borghi et al., 2007; Chao & Martin, 2000; Grèzes & Decety, 2002), which suggests that when seeing (a picture of) an object a mental simulation is also created.

According to the body-specificity hypothesis (Casasanto, 2009), these mental simulations elicited by verbal and visual information are *body-specific*, meaning that right-handers perform right-handed mental simulations of an action, while left-handers seem to perform left-handed simulations. This idea is supported by research that suggests that when imagining manual action verbs, such as *writing*, right-handers activate the left premotor cortex, whereas left-handers activate right premotor areas (Willems, Toni, Hagoort, & Casasanto, 2009; see also: Elder & Krishna, 2012; Tucker & Ellis, 1998).

Such body-specific mental simulations seem to play a role in memory and learning (e.g., Borghi, Glenberg, & Kaschak, 2004; De Nooijer, Van Gog, Paas, & Zwaan, 2013; Dijkstra, Kaschak, & Zwaan, 2007; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006). For example, right-handers have been found to remember more instructions on where to move an object when an object's handle was oriented to the right and actions had to be performed with the right hand, suggesting an affordance-based congruency benefit in memory for sequential actions (Apel, Cangelosi, Ellis, Goslin, & Fischer, 2012). In another study (Paulus, Lindemann, & Bekkering, 2009), participants had to verbally learn the functional use of objects (i.e., whether an object is used for hearing or smelling). When participants had to simultaneously perform a manual motor action (i.e., squeezing a ball) while learning these functions, performance was hindered. This was not the case when no movement or a foot movement was made during learning. This suggests that there is a crucial role for the motor system in learning the function of objects, given that hand movement was required to use the novel objects, and only executing a hand movement during learning hindered performance.

Not only can mental simulations influence learning of functional properties of objects, but it can also impact learning the meaning of novel words, which is the focus of the current study. It is not yet clear to what extent, and under what circumstances, pictures that mismatch the learner's handedness perspective can influence learning. These types of studies can provide answers to the question, when mental simulations are created and when they have an influence in cognitive processes, like word learning. One study that has investigated this issue, has shown across three experiments that right-handers' learning of object-manipulation words was hampered when presented with a picture that *mis*matched their own mental simulation (i.e., a left-handed picture) compared to one that matched their own mental simulation (i.e., a right-handed picture; De Nooijer et al., 2013). This finding was in line with the body-specificity hypothesis, which states that concepts are constituted of mental simulations, which are based on our own perceptuo-motor experiences. Therefore, seeing a picture that is incongruent with your own mental simulation of the verbal definition of the word (e.g., seeing a first-person perspective, left-handed action as a right-hander) might hinder the encoding and consolidation process of the new word. Interestingly, this effect was only found for right-handers but not for left-handers. Presumably left-handers were not hindered by a right-handed picture perspective because they come into contact with a right-handed perspective frequently: they live in a right-handed world and are, therefore, often confronted with objects meant for right-handed use, forcing them to resort to using their right hand (e.g., Stins, Kadar, & Costall, 2001). This provides left-handers with both visual and motor experience of the right-handed perspective. This differential finding for right and left-handers was in line with other research (e.g., Apel et al., 2012; Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009).

Although the research described above suggests that mental simulations are created in a large array of situations, and that they might even be necessary for many language processes, a new view of cognition, the pluralist view, argues that the creation of mental simulations might not always be necessary. According to this view cognition involves both *grounded* and *abstract* symbols (e.g., Dale, Dietrich, & Chemero, 2009; Dove, 2009; Zwaan, 2014). The emergence of this view prompts the question under what circumstances do mental simulations influence language learning? We address this question, by investigating the role of mental simulations during word learning as a function of observation perspective (i.e., observing an action from a first- or third-person perspective) and (mis)matches in handedness.

The findings of De Nooijer et al. (2013) showed that seeing first-person perspective pictures that mismatch the mental simulation created by the learner, can interfere with learning the definition of a novel verb. However, the same might not be true when pictures are shown from a third-person perspective. Research suggests that the first-person perspective is more embodied than the third-person perspective (e.g., Lorey et al., 2009). This assumption is supported by several studies that suggest that when imagining hand movements from a first-person perspective, this leads to stronger activation in motor-related structures in the brain, than when imagining an action from a third-person perspective (Lorey et al., 2009; Ruby & Decety, 2001). The first-person perspective is, therefore, more tightly coupled to the sensory-motor system than the third-person perspective, which requires additional visuospatial processing (Jackson, Meltzoff, & Decety, 2006).

Motor information not only seems to play a more important role in the first- than third-person perspective when subjects are explicitly asked to imagine hand movements, but also when they are merely viewing pictures from these different perspectives. To illustrate, when viewing images of body parts from a first-person or a third-person perspective, activity in the primary somatosensory cortex is suppressed when people view images from the third-person perspective compared to viewing the first-person perspective (Saxe, Jamal, & Powell, 2006). Motor information (as indicated by activity in the primary somatosensory cortex), therefore, seems to play a bigger role in processing pictures from a first- than from a third-person perspective. Perhaps people make, more use of motor imagery (i.e., the internal rehearsal of movements from a first-person perspective without any overt physical movement; Decety & Jeannerod, 1995; Jeannerod, 1994; Lorey et al., 2009) when viewing the first-person perspective, while they use visual imagery when viewing the third-person perspective (i.e., a third-person process involving a visual representation of an action, Lorey et al., 2009).

Support for this hypothesis comes from a study in which participants were instructed to use imagery for hand movements executed in the first-person or third-person perspective after which they had to give a laterality judgment (i.e., name the location of a single finger on the imagined hand). The participant's hands were either in a standard posture (in their lap) or in an awkward position (on their back). Under first-person imagery instructions participants responded significantly faster when they had their hands in the natural position, compared to the awkward position, suggesting that they used motor imagery for this task (i.e., having the hands in a natural “ready for action” position would result in motor facilitation and faster reaction times). For the third-person imagery the opposite pattern was found (i.e., shorter response times in the ‘hands on the back’ condition, which hinders motor imagery). These findings were taken to indicate that for first-person imagery motor imagery is used, while third-person imagery might use non-motor (visual) mechanisms (Sirigu & Duhamel, 2001).

Because mental simulations seem to be triggered automatically when an individual watches another individual performing an action (e.g., Grèzes, Frith, & Passingham, 2004; Rizzolatti & Craighero, 2004), these simulations presumably involve less motor information when observing third-person perspective actions being performed. So, if motor information is stored as part of the meaning of a word, as suggested by the body-specificity hypothesis (Casasanto, 2009) then receiving contrasting information (i.e., seeing the left-handed perspective as a right-hander) might hinder consolidation of the meaning of the word. However, if the third-person perspective does not activate motor information to the same degree, the consolidation process might not be hindered. According to this line of reasoning, right-handed participants’ learning outcomes would not be influenced by mismatching (i.e., left-handed) third-person perspective pictures (i.e., equal learning outcomes for words combined with left-handed and right-handed third-person perspective pictures), whereas it would be negatively influenced by mismatching left-handed first-person perspective pictures (i.e., lower learning outcomes for words combined with left-handed than right-handed first-person perspective pictures).

A contrasting conclusion is suggested by a study from a very different domain (i.e., motor learning), which showed that observation of same-handed models resulted in improved learning compared to observation of opposite-handed models, *regardless* of whether the model was shown from a first- or a third-person perspective (Rohbanfard & Proteau, 2011). Learning a spatiotemporal task is very different from word learning, of course, but these findings do suggest that it is possible for right-handers’ word learning to also be negatively impacted by left-handed third-person perspective pictures compared to right-handed third-person perspective pictures, which would make learning with third-person perspective pictures similar to learning with first-person perspective pictures.

To summarize, it is an open question whether word learning will be hampered when left-handed *third*-person (instead of *first*-person) perspective pictures are shown that mismatch right-handed learners' body-specific mental simulations. Therefore, Experiment 1 examines whether *third*-person perspective pictures that mismatch right-handers' hand perspective would influence their word learning to the same degree as *first*-person perspective pictures did in the De Nooijer et al. (2013) study. If the third-person perspective is indeed less embodied than the first-person perspective, and visual imagery is used here and not motor imagery then the left-handed perspective might not hamper learning compared to a right-handed and a bimanual perspective. However, if seeing the left-handed third-person perspective has the same effect in word learning as it does in a motor learning task and a mismatch in mental simulations is experienced, this perspective might hinder word learning in the same way it did when viewing first-person perspective pictures (De Nooijer, et al., 2013).

Experiment 1

Method

Participants. Participants were 227 adults (121 female) with a mean age of 34.4 years (*SD* = 11.7), recruited via Amazon's Mechanical Turk (for a description and evaluation of using Mechanical Turk in psychology experiments see Buhrmester, Kwang, & Gosling, 2011; Zwaan & Pecher, 2012). All participants were residents of the United States of America and they were not expert (master) users of Mechanical Turk. Participants were compensated with $0.75 for their participation, which required approximately 15-20 minutes.

Materials

Words. Materials consisted of 18 words from an artificial language called Vimmi (Macedonia & Knösche, 2011). An artificial language was used to avoid any influence of participants' possible prior knowledge of the target language and any possible cognates that might exist in the two languages, which might influence single word learning. Words were matched for length and were randomly coupled with the definition of a manual object-manipulation verb. For instance, the artificial word 'luko' was coupled with the definition of 'to pour' which means 'to dispense from a container'. For all words we created a picture that matched the meaning of the word. Twelve words denoted single-handed actions, and were coupled with pictures showing one hand that was executing the action. The pictures were taken from a third-person perspective and were afterwards mirrored horizontally to create an otherwise identical right- and left-handed perspective. Six words denoted bimanual actions, and therefore both hands were shown in the third-person perspective pictures. We chose to use 18 words, in order to conduct an experiment in which three blocks of six words could be used. Based on earlier experiments, we found that learning six words at a time (two words in each condition) leads to neither ceiling nor floor effects and would therefore suit our purposes. For an example of the stimuli see Figure 1.

Test. The recall test required participants to provide an English definition of the Vimmi verbs as literally as possible. During the test phase, the participants were only presented with the word from the artificial language in written form; they were not shown the pictures again.

Procedure

Participants learned the 18 words in three blocks of six words, with all conditions present in each block (i.e., 2 left, 2 right, 2 bimanual pictures), and took a test after each block. Words were counterbalanced over the three blocks, so that all words appeared in all blocks to avoid order effects. Left- and right-handed perspective pictures were also counterbalanced, resulting in the use of six lists for this experiment.

In each block, participants heard a definition of the words they had to learn, during which the word was also presented on the screen. The definition was presented twice after which the program automatically proceeded to the next word. During the second presentation of the definition of the word a picture was shown in which the denoted action was being executed. The picture had either a left, right or bimanual perspective. After six words, participants were tested on their knowledge of the words. Participants were instructed to: "Listen carefully, because after six words you will be tested on your knowledge of these words". During the post-test, participants were asked to write down the definition of the words they had just heard.

Word	**Meaning**	
Repo (Autograph)	*To write down your signature*	
	Left-handed perspective	Right-handed perspective
First-person perspective (Study 2)		
Third-person perspective (Study1 & 2)		
Kori (Tie)	*To make a knot or bow*	
Bimanual (Study 1)		

Figure 1. Example of the materials for the first- and third-person perspective conditions for Study 1 and 2a & 2b combined.

Handedness and demographic data were collected at the end of the experiment, because we did not want to provide cues about the goal of the study. Before posing the handedness question we asked participants what they thought the aim of the study was. None of the participants correctly guessed our hypothesis. Nonnative speakers of English (5%), left-handers (11%), participants who indicated that they were in a highly noisy and distracting environment while participating (i.e., people that scored a 7 or higher on a 9-point scale for background noise and distraction (1%) and participants who participated twice (2%) were removed from further analyses. This resulted in a total of 19% of participants being excluded from analyses, leaving 181 participants.

Scoring

The recall test was scored based on a straightforward coding scheme that had been developed and tested in a prior study (De Nooijer et al., 2013) in which it showed high inter-rater reliability (kappa = .82). One point was awarded for each correct answer, and a half point was given when one element of the definition was missing (e.g., 'make clothing' instead of 'to make clothing using needles for 'to knit').

Results

Data were analyzed with a Repeated-Measures ANOVA with hand perspective (left-handed, right-handed, bimanual) as within-subjects factor. Mean scores are depicted in Figure 2. There was no significant main effect of hand perspective, $F(2, 179) = .32$, $p = .726$, $\eta_p^2 < .01$.

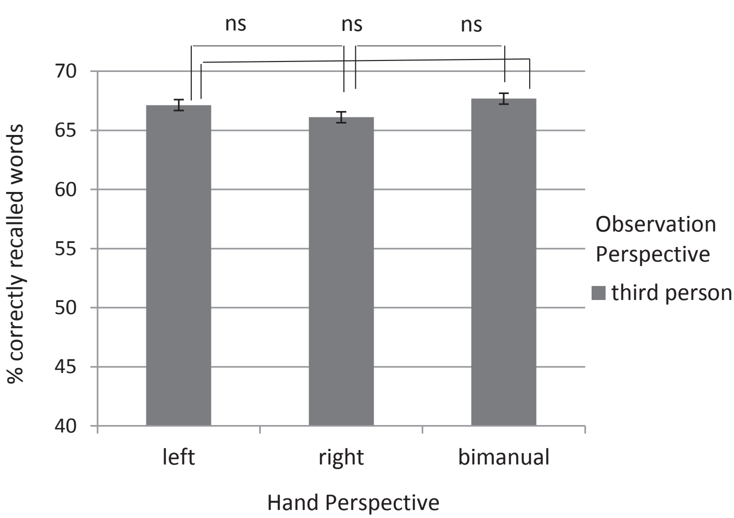


Figure 2. Percentage correctly recalled words after seeing a third-person picture perspective with a left-handed, right-handed or bimanual perspective picture (Experiment 1). Error bars indicate 95% confidence intervals.

Discussion

Experiment 1 showed that in contrast to findings with first-person perspective pictures (De Nooijer et al., 2013), learning words with third-person perspective pictures was not affected by the hand perspective shown in the picture; there was no significant difference in the percentage of definitions recalled for words learned with a left-handed, right-handed or bimanual perspective picture. This might suggest that participants did not experience a mismatch with their mental simulation of the verbal definition when presented with third-person left-handed perspective pictures. However, given that this was a single experiment, a replication would be worthwhile, which we undertook in Experiment 2a.

Moreover, it is still unclear whether presenting pictures during learning improves performance at all and therefore a no-picture control condition was added to Experiment 2a. Because it would also be helpful to establish this with regard to first-person perspective pictures, as it would help to make more sense of the findings of De Nooijer et al. (2013), we conducted a similar Experiment (2b) with first-person perspective pictures. This would not only allow for replicating their findings but would also help clarify whether left-handed first-person perspective pictures just do not contribute to learning, or whether they actually hamper learning compared to a no-picture control condition.

There are several studies that have shown an improvement in word learning when words were presented simultaneously with pictures illustrating the meaning of words from a variety of word classes (i.e., both verbs, nouns and adjectives; e.g., McGregor et al., 2007; Smith et al., 1987; Valeri-Gold, 1994). However, the simultaneous presentation of words and pictures does not always improve performance. For example, Acha (2009) found that word learning performance was best when novel nouns were learned with only a written translation of the word, compared to presenting both text and picture.

Based on the literature and the findings from Experiment 1, we can formulate several hypotheses. Firstly, we expect to replicate the finding from Experiment 1, showing that hand perspective does not affect learning with *third*-person perspective pictures (Experiment 2a). In addition, corresponding to previous findings (De Nooijer et al., 2013), we expect a main effect of hand perspective (left-handed, right-handed) for the *first*-person perspective pictures

(Experiment 2b), showing that right-handers will be negatively impacted by seeing a left-handed compared to a right-handed perspective during learning. Lastly, there are two open questions regarding the effect of pictures on word learning. Supposing that the findings of Experiment 1 are replicated, would third-person perspective pictures improve learning outcomes compared to learning without pictures (Experiment 2a)? In addition, supposing that we replicate the findings by De Nooijer et al. (2013), would right-handed first-person perspective pictures improve learning outcomes compared to learning without pictures (Experiment 2b)?

Experiment 2a & 2b

Method

Participants and Design. Participants were 162 adults in Experiment 2a (69 female) with a mean age of 34.07 years (SD =10.82), and 158 adults (76 female) with a mean age of 33.81 years (SD = 12.01) in Experiment 2b, none of whom had participated in Experiment 1. All participants were recruited via Amazon's Mechanical Turk. They were residents of the United States of America and they were not expert (master) users of Mechanical Turk, meaning that they did not come from the select pool of Mechanical Turk users that have previously shown high accuracy on a number of tasks. They were compensated with $1.00 for their participation, which required approximately 20 minutes.

The two experiments were run separately and simultaneously and were identical except for the fact that in Experiment 2a third-person perspective pictures were used, while in Experiment 2b, we used first-person perspective pictures. In both experiments there was one within-subjects factor: hand perspective with three levels (no picture, left-handed, right-handed perspective).

Materials

Words. Materials were largely the same as in Experiment 1, except that: (1) instead of six bimanual action words and pictures, six novel unimanual action words without pictures were used, resulting in six verbal definitions being learned with a left-handed, six with a right-handed perspective picture and six without a picture; and (2) pictures from both a first- and a third-person perspective were used (in Experiment 2a and 2b respectively). For an example of the stimuli see Figure 1.

Test. The tests were the same as in Experiment 1.

Procedure

The procedure was also largely the same as in Experiment 1. Participants heard the definition of the words twice, like in Experiment 1, and during the second auditory presentation of the definition they either saw no picture, or they learned the words with a left- or right-handed picture perspective (either first- or third person perspective) in the other two blocks. The block without pictures was always presented as first or last block, because we did not want to confuse the participants by randomly presenting words with or without pictures (i.e., in case we would find that learning without pictures resulted in lower learning, this might be due to participants wondering about the absence of a picture instead of the manipulation itself). Moreover, in contrast to Experiment 1, we now rotated the artificial words over definitions, to control for possible item effects (if for example in list 1, the artificial word “otu” was coupled to the definition of ‘to pour’, then ‘otu’ would be coupled to the definition of another verb in list 2, 3, etc.). This method resulted in eight lists for the first-person perspective experiment and eight for the third-person perspective experiment.

Handedness and demographic data were again collected at the end of the experiment. Left-handers (as indicated by a score lower than 40 on the Edinburgh handedness inventory) were excluded from analyses (15.4 % in 2a and 15.8 % in 2b), along with non-native speakers of English (1.8% in 2a and 5.1% in 2b), participants who indicated that they were in a highly noisy and distracting environment while participating (i.e., people that scored a 7 or higher on a 9-point scale for background noise and distraction) (0.6% in 2a and 1.9% in 2b) and participants who had participated twice in the experiment (0.6% in 2a and 0.6% in 2b) were removed prior to further analyses. This resulted in a total of 18.4% of participants being excluded from analyses in Experiment 2a, leaving 132 participants, and 23.4% of participants being excluded from analyses in Experiment 2b, leaving 121 participants.

Scoring

Scoring was the same as in Experiment 1.

Results

Mean performance scores are depicted in Figure 3. Data were analyzed to answer our three questions: Can we replicate the findings from Experiment 1? Can we replicate the findings from De Nooijer et al. (2013)? What is the effect of first- and third-person perspective pictures in word learning compared to word learning without pictures? To answer our first and third question regarding third-person perspective pictures (Experiment 2a), we performed a Repeated-Measures ANOVA with within-subjects factor hand perspective (no picture, left-handed, right-handed perspective picture). There was no effect of hand perspective, $F(2, 130) = 1.33$, $p = .269$, $\eta_p^2 = .02$, which means these findings replicated our results from Experiment 1, and additionally showed that learning with pictures did not foster learning outcomes compared to learning without pictures.

To answer our second and third question regarding first-person perspective pictures (Experiment 2b), we again performed a Repeated-Measures ANOVA with within-subjects factor hand perspective. We found a significant main effect of hand perspective, $F(2, 119) = 4.55$, $p = .013$, $\eta_p^2 = .07$. A paired-sample t-test with a Bonferroni adjusted alpha level of .017 showed that performance on words learned with the left-handed perspective was significantly lower than on words learned with right-handed first-person perspective pictures, $t(120) = 2.70$, $p = .008$, $d = .17$, replicating the findings by De Nooijer et al. (2013). However, there was no significant difference between learning without pictures and learning with the left-handed picture perspective, $t(120) = .08$, $p = .939$, $d = .01$, and although the mean scores seem to suggest that right-handed pictures would be more beneficial for learning than no pictures, this difference was not statistically significant, $t(120) = 1.96$, $p = .052$, $d = .16$.

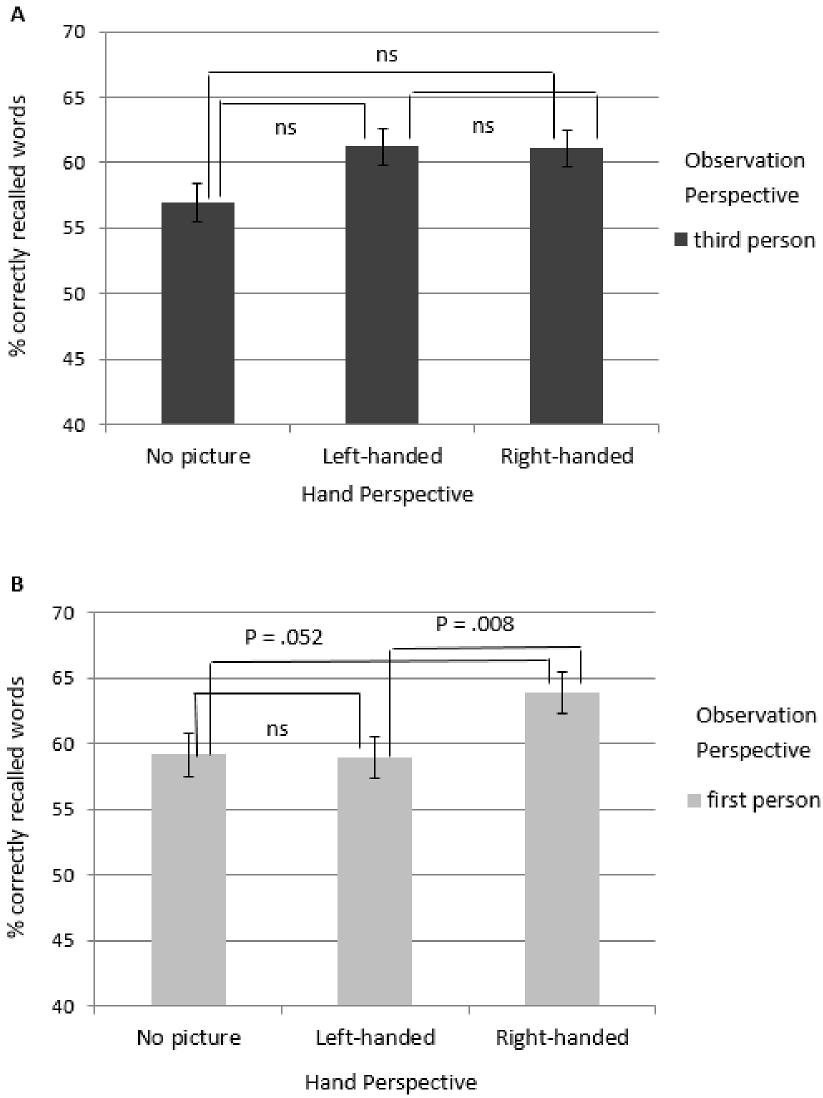


Figure 3. Percentage correctly recalled words after seeing a third-person (Experiment 2a) or first-person picture perspective (Experiment 2b) with a left-handed, right-handed picture perspective or without a picture present.

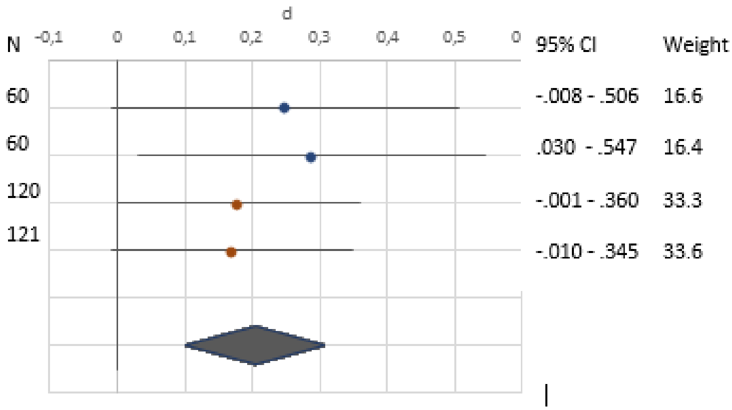
Discussion

Experiment 2a provided a replication of Experiment 1, confirming that hand perspective in third-person perspective pictures has no effect on word learning. Experiment 2b provided a replication of the study by De Nooijer et al. (2013), in that right-handers learned fewer verbal definitions when learning with a left-handed compared to a right-handed perspective picture (as in the study by De Nooijer et al., this was a small effect). We did not find evidence in either experiment for a beneficial effect of learning with left- or right-handed perspective pictures over learning without pictures. Although the results of both Experiment 2a and 2b are in line with earlier findings, we decided to perform two meta-analyses, following the guidelines and using the ESCI software of Cumming (2013), to provide more accurate estimates of the effects (or null effects) we found.

Meta-analysis of First and Third-person Picture Perspective Experiments

The first analysis concerned all available experiments investigating the effect of left- and right-handed *first*-person perspective pictures on word learning (Experiment 1, 2 and 3 from De Nooijer et al., 2013 and our Experiment 2b) and the second analysis concerned the two experiments reported here that investigated the effect of left- and right-handed third-person perspective pictures (our Experiment 1 and 2a) on word learning. Figure 4 shows a forest plot of these analyses.

For the four experiments that investigated learning with first-person perspective pictures in right-handers the analysis using the ESCI software (Cumming, 2013) showed an estimated effect size of .20 and a 95% CI that does not overlap with zero. For the two experiments that investigated third-person perspective pictures, the analysis showed an estimated effect size of .02 and a 95% CI that does overlap with zero. These meta-analyses suggest that, for right-handers, the effect that words coupled to right-handed first-person perspective pictures are better learned than words connected to left-handed first-person perspective pictures is likely to be a real effect with a small effect size, while there does not seem to be any effect for third-person perspective pictures (i.e., given the large overlap with zero, making it unlikely that these results are due to a type II error).



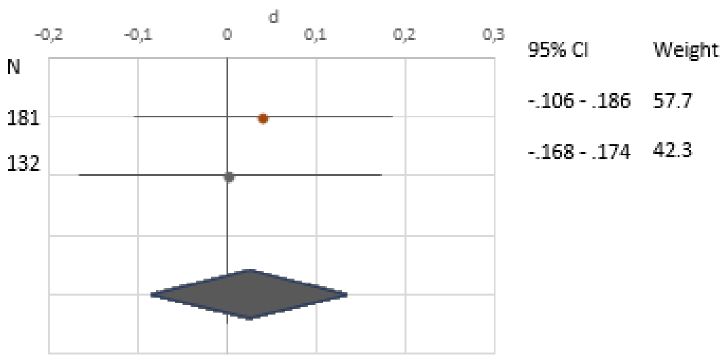


Figure 4. Meta-analysis of studies investigating the effect of first-person (top) and third-person left- and right-handed perspective pictures (bottom) on word learning

General Discussion

We investigated the effects of first- and third-person perspective pictures on word learning performance in right-handers by teaching right-handers action verbs in an artificial language with or without pictures showing a first-person or third-person left-handed or right-handed perspective. Our hypothesis that right-handers' learning would be negatively affected by mismatches between *first*-person perspective pictures and mental simulations evoked by hearing verbal definitions of action words was again supported. Right-handers presumably have strong hand-specific embodied representations, which is why they mentally simulate actions with a right-handed perspective. When shown a left-handed perspective, this mismatches the right-handers' own mental simulation, resulting in lower recall scores than for the words learned with pictures that matched the mental simulation (i.e., the right-handed perspective).

This is in line with the body-specificity hypothesis (Casasanto, 2009), which states that understanding action words requires mentally simulating one's own actions (e.g., right-handers simulate the action 'writing' with their right hand). These findings also provide evidence for the idea that memory content is associated with sensory and motor representations of an action and that the mental simulation of an action is part of its representation (e.g., Barsalou, 2008). Several studies have found a beneficial effect of seeing a picture of an object in an orientation that is congruent with a right-handers' mental simulation (i.e., a cup with the handle towards the right vs. the left) in memory tasks (Apel et al., 2012), and perceived distance tasks (Linkenauger et al., 2009). In a similar vein, right-handed participants performed better on a knot-tying task when watching a right-handed model from a first-person perspective than when watching a left-handed model (Michel & Harkins, 1985; See also: Osman, Bird, & Heyes, 2005). Watching an opposite-handed model might require some sort of additional processing effort because of the mismatch with the mental simulation evoked by the verbal information; when watching a same-handed model, in contrast, the mental simulation evoked by the verbal definition matches the one evoked by the picture.

For the third-person perspective there was no effect of pictures that mismatch the mental simulation (i.e., left-handed perspective pictures) on word learning. There were no differences in performance scores for learning words with left and right-handed third-person picture perspectives. This null effect is unlikely to be due to an undetected effect of hand perspective, given that a post-hoc power analysis with the program G*power (Erdfelder, Faul, & Buchner, 1996) indicated that the power of all experiments for detecting small effects was above .80, and moreover the small-scale meta-analysis also suggests there was no effect of picture perspective given the large overlap of the 95% CI with zero.

So what could explain the difference between the first and third-person perspective when it comes to learning words with left or right-handed perspective pictures? As mentioned in the introduction, it can be argued that the third-person perspective is less embodied than the first person perspective (e.g., Lorey et al., 2009). This could have led to the (automatic) use of a motor imagery strategy when viewing the first-person perspective pictures that would mismatch the motor imagery evoked by the verbal definition, while viewing the third-person perspective pictures may have triggered solely visual imagery (which would be in line with previous research; Sirigu & Duhamel, 2001) that would not necessarily affect the motor imagery evoked by the verbal definition.

The finding that learners are not influenced by viewing a left-handed third-person perspective is, however, not consistent with the previously described study by Rohbanfard and Proteau (2011) who found that observation of same-handed models led to significantly better learning than did observation of an opposite-handed model, regardless of perspective. The learning task used in their study, however, was very different from the one used in this study: participants were required to learn to execute a spatiotemporal motor task, whereas in our study no hand movement was required during the learning phase. There is a possibility that contrasting mismatching motor information received from watching the third-person perspective (i.e., seeing the third-person left-handed perspective) might still influence learning, but only when the motor information is necessary to complete the task, which was not the case in our study.

An alternative explanation for why participants were uninfluenced by the mismatching third-person left-handed perspective pictures might be that the third-person left-handed perspective induces a mirror like situation, which might facilitate mapping the left-handed perspective onto the observer's motor system (Alaerts, Heremans, Swinnen, & Wenderoth, 2009; See also, Hesse, Sparing, & Fink, 2009). Furthermore, people have a tendency to imitate from a mirrored image when viewing from a third-person perspective (Avikainen, Kulomäki, & Hari, 1999). The left-handed third-person perspective would then lead to executing the right-handed first-person perspective. However, in this case, one would expect a benefit of observing third-person left-handed pictures compared to third-person right-handed pictures, which we did not find. We therefore think it is more likely that the differences between learning novel verbs with first- or third-person perspective pictures, are due to the less embodied nature of the third-person perspective

Concerning the question of whether pictures would benefit action word learning at all, we cannot conclude that they do, based on these experiments. Research on the multimedia effect has provided some evidence that word (mostly noun) learning is improved when verbal information is combined with pictures (e.g., Chun & Plass, 1996; Jones, 2004; Kellogg & Howe, 1971; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968). However, this effect does not seem to be robust, and seems to depend on several factors, such as type of task (i.e., *free recall* tasks tend to produce a more reliable picture superiority effect than *cued recall* tasks; e.g., Durso & Johnson, 1980; Lotto & De Groot, 1998), and depth of processing of the picture (e.g., Craik & Tulving, 1975). These factors could have played a role in the lack of effect of pictures on word learning in our study. In addition, in contrast to other studies, all conditions in our study, including the baseline condition, provided learners with the information they needed for the test twice. Therefore, the pictures might have played a less pronounced role in learning than when the definition would have only been presented once with or without a picture. Finally, by showing the actions that were also verbally described, the information conveyed in the pictures might have been somewhat redundant (see research on the redundancy effect in multimedia, e.g., Moreno & Mayer, 2002). For object-manipulation words, a concrete type of word, pictures shown during the learning phase depict an object and the way it is held. This information is fairly redundant in combination with a definition and might therefore not have improved learning performance.

In conclusion, this study provides insight into the effect of mental simulation on word learning. We have shown that when seeing a first-person perspective picture of an action that mismatches a right-hander's mental simulation, this can negatively impact word learning. Third-person perspective pictures, however, did not show this negative effect. According to the body-specificity hypothesis and embodied cognition views, an effect of mismatching mental simulations on learning words combined with a third-person perspective might have been expected. This was, however, not the case. It appears, therefore, that observation perspective is a determining factor in the effect of mental simulations on learning behavior. These results can further our understanding of mental simulations and the extent to which they play a role in cognitive processes, like word learning.

▪ Chapter 7

Recalling the right words: Handedness effects on recalling verbs using left- or right-handed first-person perspective pictures

This chapter has been submitted for publication as:

De Nooijer, J. A., Van Wermeskerken, M., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *Recalling the right words: Handedness effects on recalling verbs using left- or right-handed first-person perspective pictures.* Manuscript in revision.

Abstract

It was recently shown that right-handers remembered fewer definitions of novel object-manipulation words in an artificial language (e.g., 'Luko' means 'to dispense from a container') when those were learned with a first-person left-handed perspective picture compared to a right-handed perspective picture. Presumably, this effect occurs because the left-handed perspective mismatches learners' mental simulation of the action evoked by the verbal definition of the word. In two experiments, the present study investigated whether the same effect would occur when asking participants to recall known verbs in their first language, combined with a matching (i.e., right-handed) or mismatching (i.e., left-handed) picture, and whether motor activation and an inter-stimulus interval that gives more time for mental simulation, would play a role in this effect. Results showed that when items were presented with a two-second inter-stimulus interval, right-handers recalled significantly more words presented with a right-handed perspective picture than a left-handed one. Eye tracking data revealed that during the inter-stimulus interval, looks to the spot where the items were previously presented were longer than to other areas. Initially, these looks were longer to the right-handed perspective pictures. Given the links that have been suggested between eye gaze to white space and retrieval of information (e.g., a list of previously encoded objects), this might mean that retrieval was more successful for the right-handed perspective pictures that match participants' mental simulation. These results suggest that (interference with) mental simulations based on someone's handedness can influence memory performance.

Introduction

When we read a sentence like “the girl wrote the report”, we create a mental simulation of this action, at least according to theories of embodied cognition (e.g., Barsalou, 1999; Glenberg & Kaschak, 2002; Stanfield & Zwaan, 2001). Mental simulations are “the reenactment of perceptual, motor, and introspective states acquired during interaction with the word, body, and mind’’ (Barsalou, 2008, p.618). These mental simulations might be different for left- and right-handers. For instance, when imagining performing a range of manual actions, such as ‘writing’, right-handers activate the left premotor cortex, while left-handers activate right premotor areas (Willems, Hagoort, & Casasanto, 2010), which suggests that left-handers create left-handed mental simulations while right-handers make right-handed simulations. This idea is supported by the body-specificity hypothesis (Casasanto, 2009), according to which reading or hearing about an action leads to the creation of body-specific simulations of this action. However, mental simulations are not only automatically created when hearing or reading words or sentences, but also when seeing (a picture of) an object, given that this may elicit motor resonance (Borghi et al., 2007).

These mental simulations are automatically created, and they might contribute to several cognitive processes, such as memory (Kent & Lamberts, 2008). For instance, a recent study suggested that motor activation induced by seeing pictures influenced memory performance (Apel, Cangelosi, Ellis, Goslin, & Fischer, 2012). In this study, right-handers remembered more instructions when an object’s handle was oriented towards the dominant hand of the participant, when an action also had to be performed with the dominant hand. This effect suggests that mental simulation can support sequential memory for actions, although it was not found for left-handers.

Furthermore, mental simulations based on both words and pictures seem to play a role in language tasks, such as word learning, A recent study suggests that right-handers’ word learning was negatively affected by possible mismatches between mental simulations evoked by verbal definitions of action words and pictures (De Nooijer, Van Gog, Paas, & Zwaan, 2013). More specifically, right-handers recalled fewer word definitions when they learned words in combination with a left-handed first-person perspective picture portraying an action of the congruent verb than when seeing a right-handed first-person perspective picture. This effect was only found for right-handers; left-handers were not hindered in learning novel verbs when seeing the right-handed perspective. Left-handers frequently come into contact with the right-handed perspective as they live in a world of which the majority of the population is right-handed. For this reason, they might not have very strong hand-specific embodied connections. For right-handers, however, seeing a picture that is incongruent with your mental simulation (i.e., seeing a left-handed action) might hinder the encoding and consolidation process of the new word.

Although the research described above suggests that mental simulations are automatically created in a large array of situations, and that they might even be necessary for many language processes, a new view of cognition, the pluralist view, suggests that we do not always need mental simulations. According to this view cognition involves both grounded and abstract symbols (e.g., Dale, Dietrich, & Chemero, 2009; Dove 2009; Zwaan, 2014). With the emergence of this view a discussion has started on when we need, or use, abstract or grounded symbols (Zwaan, 2014). Several factors could influence the extent to which mental representations are used. Whether we make a mental representation depends, for example, on the processing goal (Zwaan, 2014), but also on the experience we have with a certain action. When we have more experience with an action more grounded representations will be activated (Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008).

In this study we try to further the discussion on when we use grounded symbols, by following up on the findings by De Nooijer et al. (2013). More specifically, we investigate under what conditions mental simulations can influence memory. The study by De Nooijer et al. (2013) focused on learning new information in combination with pictures that (mis)matched the participants' handedness (i.e., participants had to learn the verbal definition of a word in an artificial language with left- or right-handed perspective pictures). It is, however, an open question whether simply *recalling* already *known* action verbs would also be affected by seeing a right- or left-handed picture perspective. Investigating this issue gives important information about when mental simulations can affect cognitive processes. Therefore, the present study investigates whether memorizing words in combination with pictures that (mis)match the mental simulation influences right-handers' memorization of the words. We only focused on right-handers, as left-handers were uninfluenced by a mismatching simulation (De Nooijer et al., 2013). In addition, we only used pictures from a first-person perspective. When using first-person perspective pictures, a motor imagery strategy might be used (e.g., Sirigu & Duhamel, 2001), given that the first-person perspective is more embodied than the third-person perspective (e.g., Lorey, Bischoff, Pilgramm, Stark, Munzert, & Zentgraf, 2009) and it recruits more sensory-motor activation (Jackson, Meltzoff, & Decety, 2006, See also: Jeannerod, 1995).

Lastly, we investigated two factors that might influence the effect of mental simulations on memorization: arm movement and inter-stimulus intervals. Research into these factors can provide more information on the question when mental simulations can influence our behavior. This can provide information on the nature and constraints of these simulations, which might help to develop our theories of embodied cognition.

Effects of movement on memory and learning

Several studies have shown that movement can improve memorization (e.g., Dijkstra & Kaschak, 2006; Engelkamp & Cohen, 1991). Children can remember more words from a list when imitating gestures that accompany the words than when only observing the congruent gestures (Tellier, 2005). Memory for actions also improves when actions during encoding are self-generated (i.e., the enactment effect; see review by Engelkamp, 1998; Nilsson, 2000). For example, children recalled more of an event they witnessed when they were allowed to gesture during recall than when gesturing was prohibited (Stevanoni & Salmon, 2005). Although these hand movements are meaningful, other studies have suggested that they do not need to be. For example, when being physically active (i.e., cycling) while learning new words, learners scored higher on vocabulary tests than learners who were physically passive, suggesting that mere motor activation can enhance word learning performance (Schmidt-Kassow, Kulka, Gunter, Rothermich, & Kotz, 2010).

Although in the studies described above motor activation seems to facilitate memory, one might also expect that motor activation interferes with memory. This idea comes from interactions that have been found between action and cognitive processes, such as memorization, which suggest that cognition shares resources with action (e.g., Barsalou 2008; Glenberg, 1997; Glenberg & Kaschak 2002). When participants are instructed to remember objects or actions while simultaneously performing a motor task, memory for the to-be-remembered objects/actions might be less efficient, because a motor task might interfere with creating a mental representation (e.g., Busiello, Costantini, Galati, & Committeri, 2011). Mental representations of actions might rely on the activation of motor affordances (i.e., potential use) for those actions. When the affordances cannot be activated, because the motor system is occupied, memory performance could decline. Activating the motor system during a memory task might, therefore, interfere with memory (e.g., Smyth, Pearson, & Pendleton, 1988). To illustrate, retaining information about manipulable objects (e.g., a hammer) in memory leads to premotor activation, but retaining information about non-manipulable objects (e.g., house) does not (Mecklinger, Gruenewald, Weiskopf, & Doeller, 2004), which suggests that working memory for objects is based on their affordances. However, a few recent studies do not find an effect of these affordances on working memory (e.g., Pecher, 2013; Pecher et al., 2013).

In sum, findings are mixed regarding the question of whether or not physical activity affects working memory. Therefore, we investigated the role of movement execution on a memory task in which action words are accompanied by left-handed or right-handed pictures. If performing a meaningless right-handed arm movement would interfere with or stimulate mental simulations, this would negatively or positively affect memory performance.

The influence of inter-stimulus intervals on mental simulations

In addition to action (in the form of arm movements) the speed with which items that have to be remembered are presented might influence whether mental simulations have an effect on memory. When items are presented consecutively there is no opportunity for rehearsal of the word in absence of the stimulus, whereas these processes can occur when items are presented with a short interval in between. When there is no time in between presentation of to-be-remembered items, rehearsal of information in between items is unlikely, because the upcoming item already has to be processed. During an inter-stimulus interval, in contrast, people can continue to memorize the words through rehearsal and associated mental simulation of the action. Participants might also make use of a process similar to the projections described by Zwaan (2014). This means that during the inter-stimulus interval sensorimotor processes, in this case eye movements, can be used to find the location where the item was previously located. People might then use grounded representations to project on this location (Zwaan, 2014). This process is thought to involve mental imagery. Mental imagery is the conscious use of grounded symbols (e.g., Barsalou, 2010) and is thought to re-activate previous motoric and sensory experiences (e.g., Gallese & Lakoff, 2005). In addition to mental simulation, the use of an inter-stimulus interval might give rise to mental imagery, which might facilitate retrieval of information (e.g., Marks, 1973; Pressley, 1976; Saltz & Donnenwerth-Nolan, 1981). Some evidence that such simulations of seen information occur during memory retrieval comes from eye tracking research. It has been suggested that when people look at a blank space presented in between items, they make eye movements to the locations of previously seen objects, and that this can facilitate retrieval of information from memory (e.g., Altmann, 2004; Johansson & Johansson, 2014; Laeng & Teodorescu, 2000; Tremblay, Saint-Aubin, & Jalbert, 2006). To illustrate, a recent study manipulated where people could look during the retrieval phase of an episodic memory task. The participants could either freely view the blank screen, maintain a central fixation, look inside a square that was congruent with the location where the to-be-recalled information was presented or look inside a square the was not congruent with the location of the to-be-recalled items. The eye-tracking data suggest that memory retrieval is facilitated when participants looked at the blank space in the square, where the to-be-recalled items were previously presented (Johansson & Johansson, 2014, see also: Olsen, Chiew, Buchsbaum, & Ryan, 2014).

Thus, an inter-stimulus interval provides an opportunity to retrieve the information that was presented when this information is no longer available. During this time, mental imagery can be used as a way of rehearsal of the information. When using mental imagery, this imagery is also body-specific (i.e., right-handed people imagine using a spoon with their right hand; Willems, Toni, Hagoort, & Casasanto, 2009). Mental imagery might, therefore, be easier for the right-handed participants after having seen a right-handed picture perspective given that this perspective matches their own motoric experience with the action. The mismatch with the picture perspective might hinder consolidation and encoding of information, which would be in line with previous studies (e.g., De Nooijer et al., 2013).

The present study

Two experiments were conducted on the effects of picture perspective (left or right-handed) on word recall; Experiment 2 was conducted to follow-up on interesting findings from Experiment 1 and will therefore be introduced later on. In Experiment 1 we presented participants with object-manipulation verbs (e.g., to cut) combined with a left-handed or right-handed first-person perspective picture that reflected the action. The picture-word pairs followed each other immediately (no break), were presented with a two second break in between (break), or were presented after the participant made a non-meaningful right arm movement needed to progress to the next word (movement).

Experiment 1

Based on the literature described above, three hypotheses and questions can be formulated. First, on the basis of the body-specificity hypothesis and prior findings regarding learning of new information, we predicted that right-handers should make right-handed mental simulations and should therefore be negatively affected by seeing an action performed by a left hand, resulting in lower recall scores for words associated with the left-handed perspective pictures. Second, whether memory performance would be enhanced, reduced, or unaffected by the execution of right arm movements, is an open question. Furthermore, when there is a two-second break between items as opposed to when items are presented immediately after each other, participants might not only automatically create a mental simulation, but also make use of the conscious process of mental imagery. It is expected that this might increase the performance gap that is expected between right-hander's recall of words presented with a right-handed compared to a left-handed picture perspective.

Method

Participants and design. 401 participants (196 male, 205 female) with a mean age of 33 years (*SD* = 12.6) were recruited via Amazon's Mechanical Turk (see Buhrmester, Kwang, & Gosling, 2011; Zwaan & Pecher, 2012) to participate in this study. Participants received $0.30 in return for their participation in the task, which lasted for approximately 5-7 minutes. All participants were residents of the United States of America. The number of participants used for this experiment was based on an a-priori power analysis with the program G*power (Erdfelder, Faul, & Buchner, 1996). The statistical power analysis was performed for sample size estimation, based on the data from the study by De Nooijer et al., 2013. The effect sizes in these experiments were considered small to medium (Cohen, 1977). With alpha .05 and power .80 the sample size for the within- group comparison is approximately N = 95. Under all three conditions we had at least 95 participants, therefore having enough statistical power to find an effect.

The experiment had a mixed design with picture perspective (left-handed vs. right-handed) as within-subjects factor and mode of presentation (no break, break, movement condition) as between-subjects factor. Participants were randomly assigned to conditions: no break (*n* = 134), break (*n* = 132), and movement (*n* = 135).

Materials. Materials consisted of 20 object-manipulation verbs (e.g., pouring, cutting, writing; for a complete list of the stimuli, see Appendix). Verbs had an average of 7.9 graphemes (*SD* = 1.1) and a mean logarithmic lemma frequency per million of 1.3 (*SD* = 0.8) determined by the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). For all verbs we created a picture that matched the meaning of the verb. In the picture a hand was shown that was executing the denoted action. The pictures were all taken from a first-person perspective and afterwards mirrored horizontally to create an otherwise identical right-handed and left-handed perspective. The verb was always presented underneath the picture. For an example of the stimuli see Figure 1.

Picture	 or
Object-manipulation word	**Pouring**

Figure 1. Example of the materials

Procedure. The experiment was implemented in the Qualtrics Survey environment. Picture-word pairs were presented on the screen for five seconds. Participants were instructed to look at the picture-word pairs carefully and to try to remember the information they were presented with. In the no break condition, the program automatically proceeded to the next item after five seconds. In the movement condition, the word-picture pair was also presented for 5 seconds, and participants subsequently had to make a hand movement to proceed to the next item (i.e., participants had to click in the top left right corner of the screen and then in the bottom right corner of the screen to proceed). Thus, there was a short break during the execution of this movement, during which the word-picture pair was no longer presented on the screen. Lastly, in the break condition, participants did not give a response to continue, but simply saw a fixation cross in the center of the screen for two seconds, before proceeding to the next trial. Of the 20 items, 10 words were presented with a left-handed perspective picture and 10 with a right-handed perspective picture; whether a word was presented with a left- or a right-handed picture was counterbalanced between participants, resulting in two lists per condition (i.e., six lists in total, given that there were three conditions) to which participants were randomly assigned. After presentation of the 20 items, participants were asked to type in as many of the words they remembered from those they just saw (i.e., a free recall task). Handedness and demographic data were collected at the end of the experiment, because we did not want to provide cues about the goal of the study but wanted to exclude left-handers.

Scoring and data analysis. We excluded participants who were left-handed (14.5%), nonnative speakers of English (3.2%), indicated that they were in a highly noisy and distracting environment while participating (i.e., people who scored a 7 or higher on a 9-point scale for background noise and distraction; 2.0%), participated in the experiment twice (which was possible, given that we collected data via Mturk; 6.2%), or did not participate on a laptop or PC (the picture and word could not be viewed simultaneously on a smaller screen, which could influence the results; 0.7%). These choices led to a total exclusion of 26.6% of the participants, leaving 293 right-handed participants for the analyses (i.e., 97 in the no break condition, 101 in the break condition and 95 in the movement condition). For all participants we scored how many words they correctly recalled that were associated with a right or left-handed picture perspective; each correct answer was awarded one point, resulting in a range of 0-10 on the words presented with a right-handed picture and a range of 0-10 on the words presented with a left-handed picture. Data were analyzed with a Repeated-Measures ANOVA with within-subjects factor picture perspective (left-handed or right-handed) and between-subjects factor mode of presentation (no break, break, movement).

Results

Figure 2 shows the percentage correctly recalled words combined with a left-handed or right-handed perspective picture for all three conditions. The ANOVA showed no significant main effect of picture perspective: $F(1, 290) = 2.37$, $p = .125$, $\eta_p^2 = .01$, nor a significant effect of mode of presentation: $F(2, 290) = 2.83$, $p = .061$, $\eta_p^2 = .02$. However, there was a significant interaction between mode of presentation and picture perspective: $F(2, 290) = 4.07$, $p = .018$, $\eta_p^2 = .03$. Follow-up paired t-tests with a Bonferroni adjusted alpha level of .017 showed that there was no difference between recall of words associated with a left-handed or a right-handed perspective picture in the no break condition, $t(96) = .77$, $p = .443$, $d = .09$, or in the movement condition $t(94) = .43$, $p = .665$, $d = .04$. However, there was a significant difference between recall of words associated with a left or right-handed picture for the break condition $t(100) = 3.02$, $p = .003$, $d = 0.35$, with fewer left-handed ($M = 42.7$, $SD = 18.7$) than right-handed ($M = 49.3$, $SD = 19.5$) words being recalled.

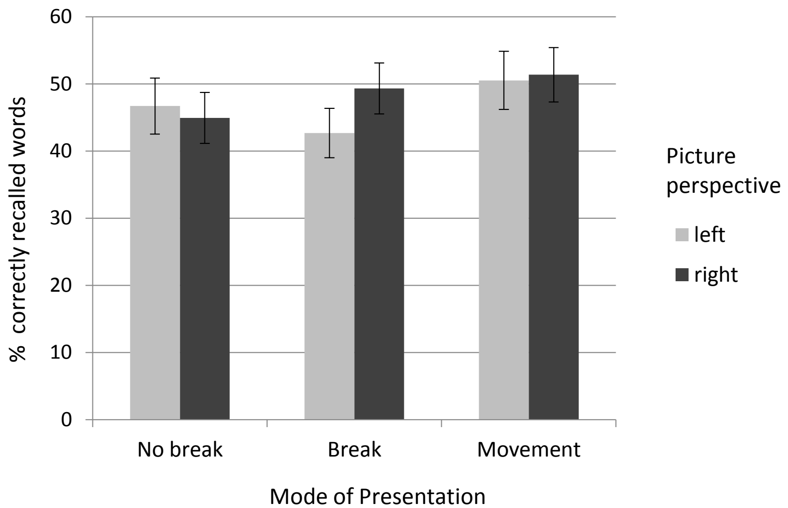


Figure 2. Mean recall performance (percentage of words correctly recalled) as a function of picture perspective and mode of presentation for Experiment 1. Error bars indicate 95% confidence intervals.

Discussion

Experiment 1 suggests that the detrimental effects of seeing a left-handed picture are not limited to learning new information, but also occur in a recall task of known action words in the subjects' first language. However, the negative effect of seeing a left-handed picture (which mismatches right-handers' mental simulation of the action) was only found when items were presented with a two-second interval, in which no actions were required during encoding. This finding suggests that the cognitive processes occurring during the two-second break in which the word and picture are no longer presented on screen, are responsible for this effect. That is, participants presumably engaged in rehearsal and mental imagery during the break in the absence of the stimuli, during which the mismatch between the left-handed picture perspective and their own right-handed mental simulation interfered with memory consolidation.

Although the movement condition also resulted in a break between items, executing the arm movement to progress to the next item might have distracted participants' attention from rehearsing the item through imagery, and consequently, picture perspective might not have affected recall in this condition. However, it should be noted that this did not hamper recall; in general, participants even seemed to score somewhat higher under the movement condition than under the break condition. Note, that our data do not exclude the possibility that having participants make a meaningful hand movement (i.e., enacting the action that had to be remembered) would have had a positive effect on recall, which seems to be suggested by several studies (e.g., Kelly, McDevitt, & Esch, 2009; Macedonia, Müller, & Friederici, 2011; So, Sim, & Low, 2012).

Because the explanation for the findings regarding the break condition is speculative, a second experiment was conducted in which eye tracking was used to study initial processing of the stimuli and to find indications of rehearsal and mental imagery during the break in the absence of the stimuli (cf. Johanson & Johanson, 2014). If we would find such indications, this would increase the plausibility of our explanation that it is the mismatch between the left-handed picture perspective and participants' own right-handed mental simulation in the absence of the stimuli that interfered with memory consolidation.

Experiment 2

In Experiment 2 we used eye tracking to investigate whether there are differences in how left-handed and right-handed pictures are initially processed or whether there are differences in continued processing (rehearsal and mental imagery) during the break in which the stimuli were absent (cf. Johansson & Johansson, 2014), which in turn could explain the difference in the number of correctly recalled words associated with the right- and left-handed picture perspective in the break condition in Experiment 1[1]. Thus , in Experiment 2 we used only the break condition from Experiment 1.

We hypothesized that if a picture that mismatches participants' own mental simulation of the action (i.e., a left-handed picture) is more difficult to process, then such pictures should (initially) capture more attention than matching, right-handed perspective pictures. However, any differences in processing *during* stimulus presentation are unlikely to explain the effect

[1] Note that we did not expect to replicate the recall performance effect of picture perspective, as this was a small effect; although we expected a similar pattern of results in Experiment 2, the main aim of this experiment was to take a closer look at the underlying mechanisms of the findings from Experiment 1.

on recall, as Experiment 1 showed that this effect occurred only in the break condition. If our theory is correct, that it is the time provided for rehearsal in this two-second break that enhances the mismatch effect between the picture and participants' own mental simulation of the action, which hampers memory consolidation, then we would expect to find indications of continued processing during the two-second break in which the stimuli are absent. More specifically, participants can be expected to look back at the exact area where the picture was previously presented (cf. Johansson & Johansson, 2014; Kent & Lamberts, 2008; see also: Brandt & Stark 1997), and the question is whether there would be any differences in looking back to the previous picture location between mismatching (left-handed) and matching (right-handed) picture items.

Method

Participants and design. Sixty-one Dutch undergraduate university students (35 male, 26 female) with a mean age of 20 years (*SD* = 1.5) participated in this study. All participants were native speakers of Dutch, had normal or corrected-to-normal vision and no neurological disorders. The experiment had one within-subjects factor: picture perspective with two levels (left-handed, right-handed picture perspective).

Apparatus and materials. Pictures were similar to those in Experiment 1. Dutch object-manipulation verbs were used, as Dutch was the participants' native language. Some items consisted of the Dutch translations of the verbs used in Experiment 1, while other verbs were exchanged for new words, because using the Dutch translations of the original verbs, led to bigger differences in the length and lexical frequency of the verbs (For a complete list of the stimuli, see Appendix). Verbs had an average of 6.4 graphemes (*SD* = 0.7) and a mean logarithmic lemma frequency per million of 1.1 (*SD* = 1.0) determined by the CELEX database (Baayen et al., 1993). Stimuli were presented using SMI Experiment Center (Version 3.3; SensoMotoric Instruments) on a monitor with a resolution of 1680 x 1050 pixels. The stimulus was either presented at the top of the screen (center of stimulus at 240 pixels) or the bottom of the screen (center of stimulus at 810 pixels) in the horizontal center of the monitor (center of stimulus at 840 pixels). Pictures were 396 x 362 pixels; the text field of the verb subtended

396 x 110 pixels. The fixation cross was 25 x 25 pixels and presented at the center of the monitor (at 840 (horizontal) and 526 (vertical) pixels). Eye movements were recorded using an SMI RED 250 eye tracker (SensoMotoric Instruments) that recorded binocularly at 250 Hz, with iView software (Version 2.8; SensoMotoric Instruments).

Procedure. Participants were seated in front of the monitor with their head positioned in a chin and forehead rest (distance to the monitor's center was approximately 70 cm) and were instructed to look at the picture-word pairs carefully and to remember the information they were presented with. At the beginning of the experiment, the eye tracking system was calibrated using a 5–point calibration followed by a 4-point validation procedure (Mean accuracy: 0.45°, *SD* = 0.20°, tracking ratio: 82.74%, *SD* = 7.66%). The word-picture stimuli were presented for 5 seconds in the middle of the screen, either at the top or bottom. Location was counterbalanced, as well as whether a word was presented with a left-handed or right-handed perspective picture, leading to the use of four lists for this experiment to which participants were randomly assigned. In between the stimuli, there was a two-second break during which a fixation cross was presented as sudden onset in the middle of the screen.

After the presentation of the 20 items, participants were presented with a free recall test, where they were asked to type in as many of the words they had just seen. Subsequently, they were requested to fill out the Edinburgh handedness Inventory (Oldfield, 1971; to avoid influencing participants during the experiment, we did not ask them about their handedness until after the experiment).

Scoring and data analysis. Eight participants (13.1%) turned out to be left-handed and could therefore not be included in the analysis, leaving data of 53 participants. Scoring of the behavioral data was the same as in Experiment 1. For analysis of the eye tracking data we created areas of interest (AoIs). During the five-second stimulus presentation AoIs were the picture (422 x 362 pixels) and the verb (422 x 110 pixels), either located at the top or at the bottom of the screen. During the two-second break, AoIs were the area where the prior stimulus had been presented (i.e., relevant AoI; 422 x 472 pixels) and the opposite area (irrelevant AoI; 422 x 472 pixels). That is, if the preceding stimulus was presented at the top, the relevant AoI would be that area, above the fixation cross, and the irrelevant AoI would be the area below the fixation cross where stimuli did appear on some of the other trials. Data from the right eye was used in the analyses.

To define fixations, we used a 40°/s velocity threshold and a minimal duration of 100 ms (in line with recommendations by Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka, & Van de Weijer, 2011). In case the deviation of calibration exceeded 1°, participants were excluded from further analyses, resulting in the inclusion of 49 participants for the eye tracking data analyses. In addition to fixations, relative dwell time was used in the analyses. Relative dwell time provides a measure of overall viewing time per AoI during the entire presentation of the stimulus, as it includes all gaze points that fall in the AoI (i.e., the sum of all gaze points within an AoI divided by the total presentation time of the stimulus). The duration of the first fixation on an AoI gives an indication of how much attention a verb or picture captured initially.

Lastly, to explore whether observers initially looked longer at the position where the stimulus was previously presented when they were looking at the white space during the two second inter-stimulus interval, we computed relative dwell time on the relevant and irrelevant area of interest (i.e., area where the stimulus was presented vs. the opposite area). We then computed mean relative dwell times for each AOI for each participant and each of the two stimulus types (picture perspective: right, left).

Results

Recall performance

Figure 3 shows the percentage correctly recalled words combined with a left-handed or right-handed perspective picture. Although the pattern of results was the same as in Experiment 1, with fewer words learned with left-handed pictures being recalled than words with right-handed pictures, as expected, a t-test for paired-samples did not reveal a statistically significant difference in the number of words recalled that were coupled with a left or right-handed picture perspective, $t(52) = 1.20$, $p = .238$, $d = 0.20$. This likely is a consequence of the lower sample size in combination with a small effect, and as mentioned before, the main aim of Experiment 2 was not to replicate the recall performance data from Experiment 1, but to provide an explanation for the data from the break condition in Experiment 1 at a process level (for which, given the nature of the eye movement data, we did have sufficient power). As such, it is sufficient that the data show the same pattern. Nevertheless, we conducted a small scale meta-analysis on the recall data from the break condition across the two experiments, following the guidelines and using the ESCI software of Cumming (2013: www.thenewstatistics.com), in order to get an estimate of the combined effect of these studies.

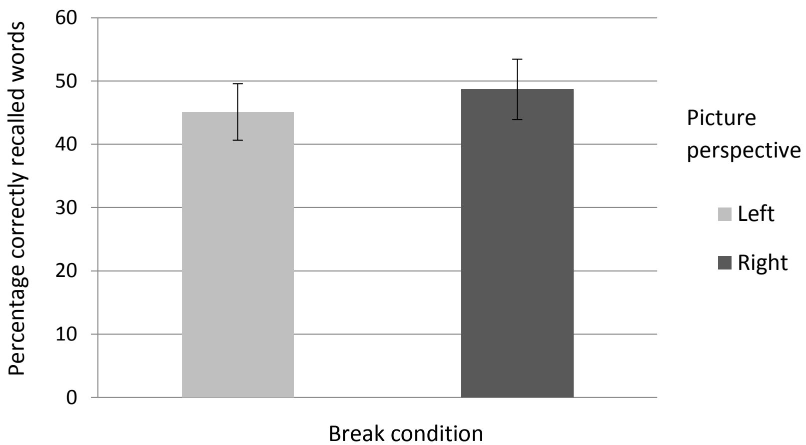


Figure 3. Mean recall performance (percentage of words correctly recalled) as a function of picture perspective and mode of presentation for Experiment 2. Error bars indicate 95% confidence intervals.

Figure 4 shows a forest plot that includes our two experiments. The point estimates, denoted by the (green) squares, indicate the difference between the mean number of words recalled associated with a right or a left-handed perspective picture (expressed in percentage correct). A positive point estimate corresponds to a numerical effect of picture perspective. Figure 4 shows that both studies yielded an effect of picture and shows the 95% CI of the combined effect in red. The combined effect shows an advantage of recall of words coupled with right-handed perspective pictures over recall of words coupled with left-handed perspective pictures. The estimated effect size in the meta-analysis is .29 and its 95% CI does not overlap with zero. This small scale meta-analysis shows that the effect that the finding that words associated with a right-handed picture perspective are better recalled than words connected to a left-handed picture perspective is likely to be a real effect, and that the reason we did not find it in Experiment 2, might be a lack of statistical power, given that the effect we were looking for is fairly small.

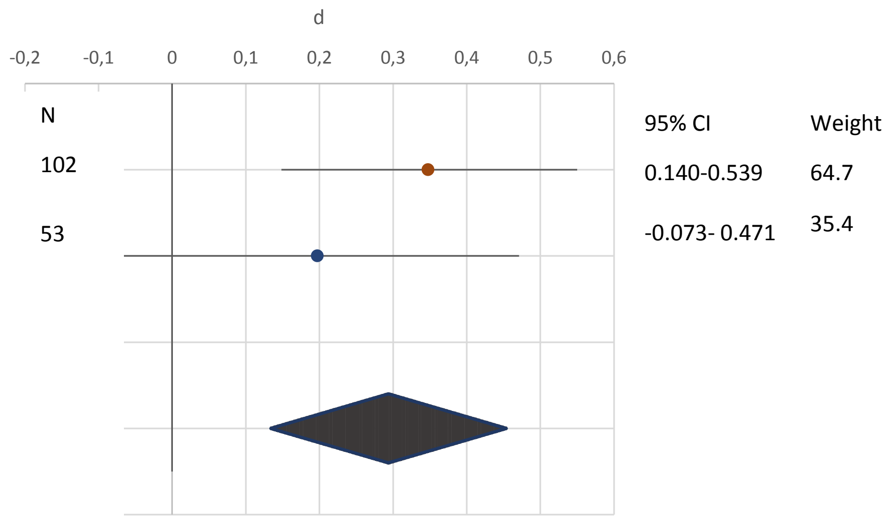


Figure 4. Results of the small scale meta-analysis.

Eye movement data

For each participant and each stimulus type (i.e., left-handed or right-handed picture perspective), mean relative dwell time and mean first fixation duration on the two areas of interest (picture, verb) were computed and entered into a 2 (Picture Perspective: left versus right) x 2 (AoI: picture versus verb) Repeated-Measures ANOVA. Means and standard deviations for the eye movement data are reported in Table 1.

Relative dwell time on stimuli. The repeated-measures ANOVA on mean relative dwell time revealed a main effect of AoI. Overall, participants were more inclined to look at the picture AoI than at the verb AoI, $F(1,48) = 413.46$, $p < .001$, $\eta_p^2 = .90$. There was neither a main effect of picture perspective, $F(1,48) = 0.12$, $p = .735$, $\eta_p^2 < .01$, nor an interaction between picture perspective and AoI, $F(1,48) = 0.17$, $p = .680$, $\eta_p^2 < .01$.

Table 1. Means (and Standard Deviations) of the relative dwell time (proportion) and first fixations (in ms) on areas of interest (AOIs) when presented with left- (left) or right-handed picture perspectives (right)

	AOI			
Measure	**Picture**		**Verb**	
Rel. dwell time	.70 (.11)		.17 (.08)	
First fixation	299.79 (187.75)		274.50 (105.04)	
	Left	*Right*	*Left*	*Right*
Rel. dwell time	.71 (.11)	.70 (.13)	.16 (.08)	.17 (.09)
First fixation	314.42 (248.01)	285.16 (154.41)	292.70 (149.57)	256.30 (106.32)

First fixation on stimuli. The repeated-measures ANOVA on mean first fixation duration did not reveal a significant main effect of AOI, $F(1,48) = 0.71$, $p = .402$, $\eta_p^2 = .02$. Although numerically, participants' first fixations on pictures with a left-handed perspective, were longer than when presented with a right-handed picture perspective (see Table 1), the effect of picture perspective was not statistically significant, $F(1,48) = 3.87$, $p = .055$, $\eta_p^2 = .08$. No interaction effect between AoI and perspective was apparent, $F(1,48) = 0.05$, $p = .826$, $\eta_p^2 < .01$.

Inter-stimulus interval. Means and standard deviations of the relative dwell time on the relevant and irrelevant AoI during the inter-stimulus interval are given in Table 2. A 2 (Picture Perspective: left, right) x 2 (AOI: relevant, irrelevant) repeated-measures ANOVA was conducted. This revealed a main effect of AoI, $F(1, 48) = 105.67$, $p < .001$, $\eta_p^2 = .69$, with participants looking more at the location where the previous word-picture pair was presented than at the opposite location (e.g., if the last pair was presented at the bottom of the screen, participants looked more towards that area on the bottom than towards the area at the top where other word-picture pairs sometimes appeared). There was no main effect of perspective, $F(1,48) = 0.45$, $p = .507$, $\eta_p^2 < .01$, nor an interaction effect between AOI and perspective, $F(1, 48) = .46$, $p = .500$, $\eta_p^2 = .01$.

It might be the case, however, that there is initially a difference in the looking patterns to the position where the words and left- or right-handed picture perspective were previously presented, but that this effect disappears after the full two seconds of the interval. To explore this possibility we conducted similar analyses over the first second of the two-second interval. This analysis did not yield a main effect of perspective, $F(1,48) = 3.69$, $p = .061$, $\eta_p^2 = .071$, but did reveal a main effect of AOI, $F(1,48) = 360.99$, $p < .001$, $\eta_p^2 = .88$, which again indicated that participants looked longer towards the location where the previous word-picture pair was presented. Interestingly, however, this time there was an interaction effect between AoI and perspective, $F(1,48) = 8.27$, $p = .006$, $\eta_p^2 = .15$. A Bonferroni corrected post-hoc test indicated that participants looked back at the location of the previous word-picture pair significantly longer when it had been a right-handed picture than a left-handed picture ($p = .004$).

The interaction between AoI and perspective is interesting in light of our hypothesis. It should be noted, however, that in the first ms of the break, when the fixation cross appears, participants may still have been looking at the location where the preceding word-picture pair was presented. Although this may explain why dwell times (at least initially) are higher for the *relevant* AOI as compared to the *irrelevant* AOI, it does not account for this interaction with perspective, because perspective was counterbalanced.

Table 2. Means (and Standard Deviations) of relative dwell time during the inter-stimulus interval where the word and (left- or right-handed perspective) picture pair was (relevant) or was not (irrelevant) previously presented.

Relevant area				**Irrelevant area**			
2 second interval		*1 second interval*		*2 second interval*		*1 second interval*	
.265 (.105)		.389 (.105)		.104 (.077)		.078 (.056)	
Left	Right	Left	Right	Left	Right	Left	Right
.261 (.108)	.270 (.113)	.375 (.097)	.404 (.119)	.105 (.076)	.103 (.088)	.081 (.057)	.074 (.067)

Discussion

Recall scores for words associated with left and right-handed picture perspectives showed the same pattern of results as in Experiment 1 (though non-significant due to smaller sample size), and a small-scale meta-analysis on data from both Experiments suggests that the overall effect that right-handers recall more words when they are associated with a right-handed than to a left-handed picture perspective does exist. An analysis of the dwell time across the entire presentation of the stimuli showed that there were no differences in looking patterns to the left- or right-handed picture perspectives that could explain these results. So overall, during the five seconds that the word and picture were presented for encoding, participants did not look at the left or right-handed picture perspective differently. However, first fixations were longer on the left-handed pictures than on the right-handed perspective pictures. This might mean that the left-handed pictures are more difficult to process (e.g., Holmqvist et al., 2011; Loftus & Mackworth, 1978; Rayner, 1998) or draw more attention because of the mismatch with participants' own mental simulation of the action. We will return to this point in the general discussion.

In terms of indications of continued processing during the two-second inter-stimulus interval, our hypothesis about people looking longer at the position where the word-picture pair had been presented was supported. Participants looked a substantial proportion of the time (on average between 25% in the entire two-second break and 39% in the first second of the break) at the position where the word-picture pairs were previously presented. Although a fixation cross was presented during the inter-stimulus interval that the participants could look at, they still looked a substantial amount of time to the spot where the items were previously presented, which might reflect processes of retrieval during this time (e.g., Johansson & Johansson, 2014).

Interestingly, during the first second of the two-second break, participants looked at the location of the previous word-picture pair longer when it was a right-handed rather than a left-handed perspective picture. This might indicate that retrieval was more successful for these right-handed perspective pictures, given that several eye-tracking studies have suggested links between eye gazes to white space and retrieval of information (e.g., Johansson & Johansson, 2014; Kent & Lamberts, 2008).

General Discussion

This study aimed to follow-up on and extend the findings from a study by De Nooijer et al. (2013), who showed that right-handers' learning of new action words (i.e., participants had to learn the verbal definition of an object-manipulation word in an artificial language), by investigating whether simply recalling known action words would also be affected by picture perspective. Experiment 1 showed that this was the case, but only when a two-second break was present in between stimuli, during which the stimulus was not shown. That is, our results suggest that remembering words associated with a picture mismatching the mental simulation generated by the observer (i.e., seeing a left-handed picture perspective as a right-hander) can indeed hinder memory for those words in a free recall task, but only when there is time for such a simulation in absence of the stimuli, this process might be complemented by the conscious process of mental imagery. These results are in line with findings by De Nooijer et al. (2013), in whose experiments there was no break between the stimuli, but verbal information was presented first without any pictorial information. During the time the verbal information was presented a mental simulation could be created that could match or mismatch the subsequently shown left- or right-handed perspective picture. It seems, therefore, that memory and learning are only influenced by mismatching mental simulations when there is time to create and repeat mental simulations, which could be complemented with more conscious processes of mental imagery.

When a right-handed arm movement was made we did not find a difference in recall scores between words associated with a left- or right-handed picture. However, when presented with a two-second interval, imagery might have been used as a retrieval strategy, which may not have been possible when the items were presented straight after each other or when a movement had to be made between the stimuli. This might be one explanation why we only found a difference between the number of words recalled associated with the left- and right-handed perspective when a two second-break is presented. A period in which the information is retrieved before going on to the next item is a prerequisite for mismatching mental simulations to have an effect on memory. The inter-stimulus interval could have made the pictures salient (because the pictorial information could be retrieved), enhancing the mismatch effect. This is in line with recent suggestions that grounded congruency effects might depend on task conditions (Lebois, Wilson-Mendenhall, & Barsalou, 2014)

The idea that a process of information retrieval takes places during the inter-stimulus interval is supported by the eye-tracking data from Experiment 2. Other eye-tracking research has suggested that when looking at a blank space in between items, people make eye movements to the location where the to-be-recalled items were previously presented. These eye movements can facilitate retrieval of information from memory (e.g., Altmann, 2004; Johansson & Johansson, 2014; Laeng & Teodorescu, 2000; Tremblay, Saint-Aubin, & Jalbert, 2006). The inter-stimulus interval provides participants with an opportunity to retrieve the information they have been presented with when the information was no longer available. If this retrieval takes place via mental imagery, then participants would look significantly longer at the exact spot where the previous word-picture pair was presented, which is supported by the data from Experiment 2. The initially longer looks to the right-handed perspective pictures during the inter-stimulus interval might reflect the more successful retrieval of the words associated with these pictures. These results are in line with other studies in the field (e.g., Johansson, Holsanova, & Holmqvist, 2005). Eye movements to previous locations have been suggested to facilitate retrieval of information (Ferreira, Apel, & Henderson, 2008), or provide a spatial index that is part of the memory representation that is remembered which might have a functional role in the mental imagery process (Laeng & Teodorescu, 2002). Looking to the spot where information was encoded when retrieving information seems to be a robust effect, as even a week after encoding, looks to areas where stimuli were presented were longer when recalling the previously encoded information (Martarelli & Mast, 2013).

In addition to eye movements during the inter-stimulus interval, we investigated the looking patterns during the time when the verbal and visual information was presented. These analyses would not be able to explain the effects found in Experiment 1, given that any differences in processing the word and picture during encoding between the left- and right-handed items would presumably have affected recall in all conditions, whereas that effect was only found in the break condition. Nevertheless, it is still interesting to investigate if there are differences in how these pictures are initially processed as it can give more insight in the underlying processes and it is possible that initial processing only affects recall in combination with a break in which rehearsal or imagery can take place. These data from the encoding phase suggest that during a first fixation people tended to look longer at the left-handed than the right-handed perspective pictures, which might indicate that the left-handed pictures were more difficult to process or attracted more attention because they mismatch participants' own handedness. This hypothesis is supported by findings that longer fixation duration reflects deeper and more effortful processing (Holmqvist et al., 2011). For example, longer fixation durations have been found when processing low frequency compared to high frequency words (Rayner, 1998), objects that are placed out-of-context versus those that fit the context (Loftus & Mackworth, 1978) and viewing face images versus natural scenes (Guo, Mahmoodi, Robertson & Young, 2006). When analyzing the complete five seconds that participants could view the word-picture pair, no difference in looking patterns between the left- and right-handed perspective pictures was found. The reason for this might lie in the length of the viewing time. Research suggests that the informativeness of the locations that are fixated within the picture decreases over time, when less informative details might be inspected (Antes, 1974). This might have happened in this study as well. People could have initially made a distinction between the left- and right-handed perspective pictures, which led to longer fixations to the left-handed pictures, but after initial processing of the left- or right-handed action, attention could have lingered elsewhere and looking patterns equalized for the two types of pictures.

In conclusion, although the effect of mental simulations on learning and memory tasks is small, this study and others in the field (e.g., Apel et al., 2012; De Nooijer et al., 2013) suggest that we do make mental simulations after hearing or seeing concrete actions being described or depicted, and that mismatches between our body-specific mental simulation and visual information cannot only hinder learning but also memory performance, at least for right-handers. There are, however, restrictions to when we use these grounded symbols as we only found an effect of mismatching handedness perspective when items were presented with an inter-stimulus interval. This seems to be in line with the pluralist view of cognition (e.g., Dale et al., 2009; Dove, 2009), as it appears that grounded symbols only affect cognitive processes such as memory in certain circumstances. These results can help further build our understanding of mental simulations.

Appendix

English verbs (exp 1)	Dutch verbs (exp 2)
1. Calling	1. Bellen (calling)
2. Throwing	2. Gooien (throwing)
3. Erasing	3. Gummen (erasing)
4. Beating	4. Kloppen (beating)
5. Cutting	5. Knippen (cutting)
6. Scraping	6. Krabben (scraping)
7. Clicking	7. Klikken (clicking)
8. Stirring	8. Roeren (stirring
9. Slicing	9. Snijden (slicing)
10. Stamping	10. Dammen (playing chess)
11. Stapling	11. Kaarten (playing cards)
12. Marking	12. Kammen (combing)
13. Shoveling	13. Lijmen (gluing)
14. Drawing	14. Persen (squeezing)
15. Scrubbing	15. Schaken (playing chess)
16. Pouring	16. Schuren (sanding)
17. Writing	17. Smeren (spreading)
18. Sprinkling	18. Verven (painting)
19. Perforating	19. Zagen (sawing)
20. Ironing	20. Zemen (wiping)

▪ Chapter 8

Picturing meaning: An ERP study on the integration of left- or right-handed first-person perspective pictures into a sentence context

This chapter has been published as:

De Nooijer, J. A., Gootjes, L., Van Gog, T., Paas, F., & Zwaan, R. A. (in press). Picturing meaning: An ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context. *Language, Cognition and Neuroscience.*

Abstract

Verbal and pictorial information are often processed together. Therefore, knowing how and when information from these modalities is integrated is important. In this ERP study we investigated integration of pictorial information into a sentence context. Right-handed participants heard sentences containing manual action verbs (e.g., "You are slicing the tomato"), while seeing a picture of a manual action. Pictures matched or mismatched the sentence content and the participants' handedness (i.e., pictures showed a left or right-handed perspective). Results showed a larger N400-amplitude for content-mismatching than for content-matching sentence-picture pairs. The N400-amplitude was not larger when the picture mismatched the participants' handedness. However, participants responded faster to right than to left-handed perspective pictures. This study suggests that with a sentence context, pictures are integrated with verbal information, but mental simulations either do not play a role in this process or this role might be too small to be visualized in the N400.

Introduction

Deriving meaning from the integration of multiple sources of information is an essential part of language comprehension. Many researchers have searched for ways in which integration of verbal elements can be investigated. In the 1980s it was first shown that the difficulty to integrate two semantic constructions, becomes manifest in the ERP waveform by a negative going peak around 400 milliseconds after a mismatch occurs (Kutas & Hillyard, 1980). For example, the word “cry” in the sentence “the pizza is too hot to cry”, is difficult to integrate with the rest of the sentence, which results in an N400-effect with a centro-parietal maximum. Although there is some debate as to what stage in language processing the N400 reflects, lexical selection (i.e., word-level) or lexical integration (i.e., discourse-level), the account of lexical integration seems to be favored (Salisbury, 2004). According to this account the N400 is an index of semantic processing that reflects the neural mechanisms of semantic integration into a context (Brown & Hagoort, 1993). Initially, studies on this semantic integration and the N400 have investigated integration of *verbal* elements into a sentence context (e.g., Connolly & Phillips, 1994; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kutas & Hillyard, 1980) or broader discourse (e.g., Nieuwland & Van Berkum, 2006; Salmon & Pratt, 2002; Van Berkum, Hagoort, & Brown, 1999). More recently, however, researchers have started to focus on the question of how information from other modalities, such as visual information (e.g., pictures), is integrated into a verbal context. It is very common to simultaneously encounter pictorial and verbal information, for example when reading a magazine or browsing the internet. Therefore, an interesting question is how information from these two sources is integrated. Concerning this type of integration, it was suggested that the language system incorporates semantic information coming from linguistic and extralinguistic domains over a similar neural time course and by recruitment of overlapping brain areas (Willems, Özyürek, & Hagoort, 2008). Thus, when we refer to integration of verbal and pictorial information, we mean the processing of information from two modalities into one coherent whole. This view of how integration works is in line with the one-step model of language comprehension. According to this model every source of information, whether linguistic or extralinguistic, immediately (i.e., as soon as the information from two modalities becomes available) constrains the interpretation of an utterance (e.g., Hagoort & Van Berkum, 2007; Tanenhaus & Trueswell, 1995). The visual (pictorial) context can, therefore, influence word recognition at the earliest moments during language processing (e.g., Spivey-Knowlton, & Sedivy, 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

This integration of verbal and pictorial information was investigated at the word level. For example, when participants were first presented with a verbal object name at the categorical (e.g., dog) or specific (e.g., collie) level and then with a black and white picture that either matched or mismatched the object name, an N400-effect was found for both a basic and a subordinate mismatch (Hamm, Johnson, & Kirk, 2002). In a similar vein, when acoustically presenting 19-month olds and adults with words at the categorical level that were either congruent or incongruent with a picture content, both groups showed an N400-effect (Friedrich & Friederici, 2004; see also: D'Arcy & Connolly, 1999). However, a mismatch between verbal and pictorial information not only leads to difficulties with the integration of information at the word level but such difficulties might also occur at the sentence or discourse level. Several studies have investigated the integration of pictorial information into a sentence context. In some studies on sentence-picture integration, pictures were not presented simultaneously with the sentence, but with the picture as a replacement of a word in the sentence (e.g., Ganis, Kutas, & Sereno, 1996; Nigam, Hoffman, & Simons, 1992), where an N400-effect was found when a sentence was completed with a context inappropriate picture, implying that when the picture was congruent, the information from both modalities could be integrated (For similar results see: Knoeferle, Urbach, & Kutas, 2011; Wassenaar & Hagoort, 2007). Behavioral results have also shown that replacing words with pictures does not disrupt sentence processing (Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986). However, also when the pictures only offer additional information, for example when participants were presented with short stories while simultaneously seeing a line drawing of an object that is congruent or incongruent with the context, an N400-effect for congruency was found (Willems et al., 2008). The finding that the N400-amplitude was larger for the incongruent items suggests that semantic integration was more difficult for the incongruent than for the congruent items. This shows that people even attempt to integrate information gleaned from pictures that offer additional information into the larger sentential context.

In addition to the study of integration of verbal and pictorial information, several studies investigated the integration of two sources of pictorial information. Interestingly, even without the presentation of verbal information, pictures themselves can lead to mismatches, resulting in integration difficulties. When the content of pictures is incongruent with world knowledge, an N400-effect arises, for example, when presenting participants with colored pictures of people performing simple incongruent actions, (e.g., a woman cutting bread with a saw; Proverbio & Riva, 2009), tools shown in a false orientation (e.g., a screwdriver held horizontally where the screw is shown vertically; Bach, Gunter, Knoblich, Prinz, & Friederici, 2009), incorrect tool use (e.g., using a screwdriver to open a lock; Bach et al., 2009), inappropriate passing-receiving pictures (e.g., seeing a hand in a grasping position, for both the passing and receiving action; Shibata, Gyoba, & Suzuki, 2009) or when stories in the form of a series of pictures have an incongruent picture ending (West & Holcomb, 2002; see also: Mudrik, Lamy, & Deouell 2010). These action-elicited N400 waves resemble the shape and timing of linguistic N400 waves, suggesting that the same neural mechanisms are involved in linguistic integration, as well as integration of information from the visual modality (Amoruso et al., 2013).

The research described above seems to suggest that verbal information and visual information can easily be integrated with one another, on a word, sentence or discourse level and even when two sources of information from the visual modality have to be integrated with each other, this is easily done. Even though language and sensorimotor processes seem to be integrated during the comprehension of everyday actions, it is still unclear how this happens in the brain (Amoruso et al., 2103). One factor that might be involved in the integration of information is the mental simulation created when processing information, as is stipulated in theories of embodied, or grounded, cognition. It has been suggested that the N400 can be understood within an embodied cognition framework (e.g., Hald, Marshall, Janssen, & Garnham, 2011), meaning that motor information can modulate meaning-related processes indicated by the N400 (Amoruso et al., 2013). Mental simulations also elicit (pre)motor activity (e.g., Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006) and might therefore influence the integration process. These mental simulations are evoked by both verbal and visual information. Studies have shown that motoric information is automatically activated when seeing an object or a picture of an object (Borghi et al., 2007; Sumner et al., 2007). Also, when reading an action word, the motor system becomes active (Hauk, Johnsrude, & Pulvermüller, 2004). Reading and seeing the same (manual) action should, therefore, in part lead to the same activation patterns in the brain which might facilitate integration.

If the mental simulation, that is, “the reenactment of perceptual, motor, and introspective states acquired during interaction with the world, body, and mind.” (Barsalou, 2008, p. 618), evoked by the verbal information and the picture are congruent, this might facilitate the integration of verbal and pictorial information. However, pictures mismatching the mental simulation might hinder the integration of the picture in the sentence context. According to the body-specificity hypothesis hearing about an action leads to the creation of a body-specific mental simulation of that action. For example, when hearing the sentence: “You are stirring in the pot” right-handers would make a right-handed mental simulation (Casasanto, 2009). Therefore, seeing a left-handed picture perspective might provide a mismatch with the mental simulation and hinder integration of the picture in the sentence context, which could be reflected in the N400-component. Behavioral studies have shown that both memory and learning can be hindered when a perspective shown mismatched the participants’ perspective. For instance, a recent study showed that motor activation induced by seeing pictures influenced memory performance (Apel, Cangelosi, Ellis, Goslin, & Fischer, 2012). Right-handers could remember more instructions when an object’s handle was oriented to the right and actions also had to be performed with the right hand. Also, research has shown that when learning new words coupled to a picture of a left or a right-handed picture perspective, right-handers recall fewer word definitions, when the picture seen during learning mismatches the right-handed mental simulation evoked by the verbal definition (De Nooijer, Van Gog, Paas, & Zwaan, 2013). Behavioral results can, however, reflect later processes than those reported from ERP data. If the effects of handedness on memory and learning tasks as reported above result from early processing, then a mismatch between the (right-handed) mental simulation evoked by the action verb and the left-handed picture will be reflected in the N400. If it is the result of slower-acting processes, it might be reflected in longer post-sentential reaction times to these items.

To summarize, we asked two questions. First, can we replicate the finding that pictorial information can be integrated with verbal information that is conveyed in a single sentence? We hypothesized on the basis of the current literature that pictures that mismatch the content of the sentence would evoke a larger N400-amplitude than pictures that match the sentence content, given that the mental simulations created by verbal and pictorial information are integrated with each other into a single sentence context.

Second, if pictorial information can be integrated with verbal information, is integration then facilitated when the handedness perspective of the picture matches that of the participant? Here we hypothesized that the N400-effect would be modulated by the hand perspective, where the largest N400 should occur for sentence-picture pairs that mismatch both in content and hand perspective (i.e., a left-handed picture). When both content information and handedness perspective mismatch, the integration process might be hindered. This effect might be strengthened by using sentences that are formulated in the second person perspective, such as "you are stirring in the pot", (e.g., Sato & Bergen, 2013), combined with using first-person picture perspectives, which prime an actor's perspective.

Finally, on a sentence-picture verification task, longer response times are expected on the left-handed perspective items. In line with previous research (e.g., Apel et al., 2012; De Nooijer et al., 2013) we only expected this effect for right-handers, which is why we conducted the EEG experiment only with right-handers. To foreshadow, we did not find a perspective effect on ERPs, only on behavioral data. Based on a suggestion by an anonymous reviewer, however, we decided to test whether indeed only right-handers were influenced by handedness perspective in our sentence-verification task (which is different from tasks used in prior research on handedness effects). Therefore, this task was subsequently also investigated with left-handers, for whom we did not expect to find any differences in reaction times as a function of left- or right-handed perspective pictures (e.g., Apel et al., 2012; De Nooijer et al., 2013).

By investigating these questions, this study might contribute to answering the larger question of how language and sensorimotor processes are integrated during the comprehension of everyday actions, given that it is still unclear how this happens in the brain (Amoruso et al., 2013). With this study, we try to unravel one factor that might be relevant in this issue, namely the creation of mental simulations, by investigating whether pictures that mismatch the viewer's mental simulation hinder the integration process. Moreover, to move theories on grounded cognition forward, it is necessary to focus more on when we need or use grounded symbols and mental simulations and to try to understand the nature of these mental simulations (Dale, Dietrich, & Chemero, 2009; Dove 2009; Zwaan, 2014). This study might contribute to answering such questions, as it can provide insight into whether mental simulations influence the integration of verbal and pictorial information, on which no information is available thus far.

Method

Participants

Twenty-five (16 female) undergraduate psychology students with a mean age of 23.2 years (*SD* = 3.2) participated in this study as part of an EEG tutorial. Participants were naïve to the experimental questions. The experiment lasted approximately 10 to 15 minutes. All participants were native speakers of Dutch, had normal or corrected-to-normal vision and no neurological disorders. To avoid influencing participants during the experiment, we did not ask them about their handedness until after the experiment. One participant turned out to be left-handed and could therefore not be included in the analysis, leaving 24 participants. Data of another 3 participants were not of good enough quality (i.e., fewer than 20 segments remained in each condition after artifact rejection) leaving 21 participants for the final analyses.

In addition, 25 left-handers (16 female), of whom 24 were confirmed to be left-handers according to the Edinburgh Handedness inventory (meaning that one was ambidextrous and was therefore not included in the sample), participated in the sentence-picture verification task. These 24 participants had a mean age of 20.5 years (*SD* = 2.0) and received either course credit or a small monetary reward for participation.

Materials

Material consisted of 40 Dutch sentences and pictures. Sentences always consisted of the same four elements; the second-person pronoun 'you' followed by a manual action verb, a definite article and an object (e.g., 'Jij snijdt de tomaat' meaning '*You are slicing the tomato').* Given that Dutch is an SVO language, the verb always appeared in the second position. Verbs had an average of 5.1 phonemes (*SD* = 1.39) and a mean log frequency of 1.0 (*SD* = 0.7). All sentences were recorded in a sound-attenuated room, spoken at a normal rate by a native Dutch female speaker. For all sentences we created a picture that matched the meaning of the sentence. In the picture a hand was shown that was executing the denoted action. The pictures were all taken from a first-person perspective and afterwards mirrored horizontally to create an otherwise identical right- and left-handed perspective. In addition, we made sure that all the hands depicted in the pictures were as neutral as possible (no rings, long nails, nail polish, etc.). The hands could, therefore, not be easily identified as being either male or female. This was done because handedness effects could possibly be greater in cases where the depicted hand has greater resemblance to the participant's hand. For an example of the stimuli see Figure 1.

Design and Procedure

In a within-subjects design, participants were presented with spoken sentences combined with pictures that could (mis)match the action verb in two ways: (1) the content of the verb matched the content of the picture (i.e., hearing 'you are slicing the tomato' and seeing a hand with a knife 'slicing' a tomato) or mismatched the content (i.e., hearing 'you are slicing the tomato' but seeing a hand 'writing' something). (2) The hand perspective could match or mismatch the right-handed participants' hand perspective (i.e., seeing a right-hand slicing a tomato, vs. a left-hand slicing a tomato). This resulted in four conditions: (1) content and hand perspective match, (2) content match, hand perspective mismatch, (3) content mismatch, hand perspective match, (4) content and hand perspective mismatch.

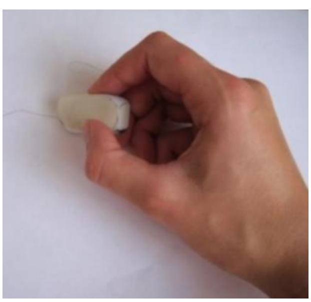
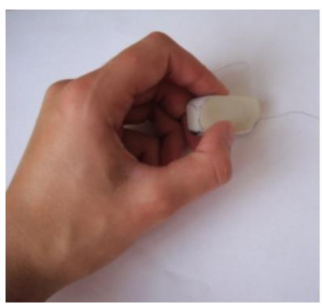
Sentence: “ You are stirring in the pot”	
Content match + hand perspective match	*Content match + hand perspective mismatch*
	
Content mismatch + hand perspective match	*Content mismatch + hand perspective mismatch*
	

Figure 1. Example of the materials for four conditions

During the EEG experiment, which was implemented in E-Prime (Schneider, Eschman, & Zuccolotto, 2002) participants were seated in a comfortable chair and were instructed to minimize movement. They heard the sentences in four blocks, with each sentence being presented in each block, but coupled with a different picture, so that each sentence appeared once in each condition (This procedure is similar to the one used in e.g., Bach et al., 2009; Friedrich & Friederici, 2004). To prevent participants from guessing what the mismatching picture would be, we made sure picture-sentence pairs were never exchanged (e.g., if the mismatching picture for the sentence: “You are writing the report” was “cutting a tomato”, then the mismatching picture for the sentence “You are cutting the tomato” would not be “writing a report”). Blocks were counterbalanced across lists, resulting in the use of four lists. Sentences were presented in a randomized order within blocks. Pictures were displayed for two seconds at the onset of the verb. All pictures were presented in the center of the screen and were 400 x 400 pixels. The audio file lasted for three seconds after which the question: “Does the picture match the content of the sentence?” appeared. Participants were instructed to indicate as fast as possible whether the content of the sentence matched the picture by pressing the ‘J’ (for ‘ja’ = yes) or ‘N’ (for ‘nee’ = no) button on the keyboard with their dominant hand (the right or left middle finger and the right or left index finger respectively). The ‘J’ and ‘N’ are placed above each other on the keyboard, which should prevent an interference effect of button placement with the left and right-handed perspective pictures. Reaction times were calculated counting from the end of the sentence. Triggers were placed in order to mark the audio onset of the verb. Before the start of the experiment two practice trials were presented which contained different critical words than used in the main part of the experiment. Participants were told to attentively listen to and watch the stimuli and to blink their eyes only in between the sentences when a fixation cross was presented for 500 ms in the center of the screen, at the beginning of each trial.

EEG recording

EEG was recorded from 32 electrode sites across the scalp using active Ag/AgCl electrodes (BioSemi, Amsterdam, the Netherlands, ActiveTwo amplifier system) placed in an elastic cap according to the 10/20 system. Electrodes were placed on standard sites (Fz, Cz, Pz, Oz, AF3, AF4, Fp1, Fp2, F3, F4, F7, F8, FC1, FC2, FC5, FC6, T7, T8, C3, C4, CP1, CP2, CP5, CP6, P3, P4, P7, P8, O1, O2, PO3, PO4). Eye movements and blinks were monitored by four additional electrodes. Horizontal eye movements (HEOG) were measured by placing electrodes on the left and right outer canthi of the eye. Vertical eye movements (VEOG) were measured with electrodes above and below the left eye. Lastly, reference electrodes were placed on the left and right mastoid. Recordings were amplified using an ActiveTwo amplifier system and sampled at 512 Hz.

EEG analysis

EEG data were re-referenced offline to the linked mastoids. Segments of 1100 ms were created, including 100 ms before the target word onset (i.e., the action verb). Epochs were filtered with a 0.01–40 Hz band filter, and corrected for eye movements using the algorithm of Gratton, Coles, and Donchin (1983). Segments were only analyzed when the correct answer was provided by the participant. In each condition the maximum number of data segments per participant was 40. If fewer than 20 segments remained in each condition after artifact rejection the participant was excluded from further analysis (3 participants), leaving 21 participants. From the remaining 21 participants on average 9 % of the segments were rejected because of artifacts (12 % in the content match, hand perspective mismatch, 12 % in the content and hand perspective match, 7 % in the content mismatch, hand perspective match and 6 % in the content and hand perspective mismatch condition). Segments were normalized on the basis of the 100 ms pre-stimulus baseline. ERPs were calculated for each participant by averaging trials for each electrode and condition separately.

We used two time windows, an N400 time window (300-550 ms) and a late time window (600-900 ms). A late time window was included because some studies have found a broad negativity following the N400 in a similar design (e.g., Mudrik et al., 2010). We, therefore, included this time window, to investigate whether our results were in line with such a finding. On the basis of previous studies (e.g., Federmeier & Kutas 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010) the N400-effect for sentence-picture integration could be expected to have a more frontal maximum than the centro-parietal maximum that is usually reported (e.g., Kutas & Hillyard, 1980), however other studies found a centro-posterior N400 distribution for picture-sentence integration (e.g., Friedrich & Friederici, 2004; Knoeferle et al., 2011). To be able to detect where the N400-effect is strongest, we created four quadrants plus a midline section to examine the time windows. All recorded channels are included in these sections: left anterior (F3, F7, FC1, FC5, C3 AF3, FP1, T7), right anterior (F4, F8, FC2, FC6, C4, AF4, FP2, T8), left posterior (CP1, CP5, P3, P7, O1, PO3) right posterior (CP2, CP6, P4, P8, O2 PO4), midline (Fz, Cz, Pz, Oz).

Statistical analyses

For the behavioral data of the right-handers two Repeated-Measures ANOVAs were conducted; one based on participant variability and one based on item variability. Both analyses were conducted with content (picture-sentence [mis]match) and perspective (picture perspective [mis]match) as within-subjects factors for the reaction times on the sentence-picture verification task. Similar analyses were separately conducted for the data of the left-handers. Only reaction times to correct answers were analyzed. In addition, reaction times faster than 100 ms or slower than 3000 ms were excluded from analyses. As a result 7.7% of the reaction time data from the right-handers and 8.3% of the reaction time data from the left-handers was not analyzed.

For the EEG data of the right-handers a 5 x 2 x 2 Repeated-Measures ANOVA was conducted with region (quadrants plus midline), content and perspective as within-subject factors for both the N400 and the late time window. In addition, to explore whether there are any differences in how the handedness perspective is processed by the two hemispheres, we performed an additional analysis with hemisphere (left, right), content (congruent, incongruent) and perspective (left, right) as within-subjects factors. When assumptions of sphericity are violated, results of Multivariate analyses are reported, given that these are not limited by the sphericity assumption (Jennings, Cohen, Ruchkin, & Fridlund, 1987). Only results with a p-value < .05 are interpreted as significant. When multiple comparisons are made, results were only considered as significant when they are Bonferroni-corrected for multiple comparisons.

In addition to classical frequentist Repeated-Measures ANOVAs, we also used Bayesian Repeated-Measures ANOVAs. Bayesian statistics allows us to calculate the probability a hypothesis is true, given the data. Bayes Factors for the separate effects in the Repeated-Measures ANOVA were computed with JASP 0.6 (http://jasp-stats.org). The Bayes Factor used is the inclusion Bayes Factor (BF_{inc}), which is an average of the likelihood of models that include the effect. It, therefore, compares all models with a certain factor against all the models without that factor. BFinc can be interpreted as follows concerning the evidence for the alternative hypothesis: $-\infty < B \leq 0.1$ is considered strong evidence against, $0.1 < B \leq (1/3)$ is substantial against, $(1/3) < B < 1$ barely worth mentioning against, $1 \leq B < 3$ barely worth mentioning for, $3 \leq B << 10$ substantial for, $10 \leq B < \infty$ strong for (Jeffreys, 1961). For example, when the BFinc is .02 this means that given the data, it is 50 times more likely that the null hypothesis is true, compared to the alternative hypothesis (i.e., the inverse of the inclusion Bayes Factor ($1/BF_{inc}$)).

Results

Behavioral data

Mean accuracy and reaction times on the accurate items are given in Table 1. Because of the high accuracy scores (see Table 1) we only analyzed the reaction times. The ANOVA for the right-handers showed no significant effect of content, $F(1, 23) = .56$, $p = .462$, $\eta_p^2 = .16$, BFinc = .23, on the reaction times on the accurate items. However, there was a significant effect of perspective, $F(1, 23) = 4.46$, $p = .046$, $\eta_p^2 = .16$, BFinc = .41, indicating that participants were faster to react to the right-handed than to the left-handed pictures. No significant interaction between content and perspective was found, $F(1, 23) = .02$, $p = .881$, $\eta_p^2 = .001$, BFinc = .097. Although we did expect to find an interaction between these factors, the null hypothesis is, given these data, over 10 times more likely than the alternative hypothesis (i.e., the inverse

BF), which provides strong evidence for the null hypothesis. In addition, the item analyses showed no effect of content, $F(1, 39) < 1$, $p = .687$, $\eta_p^2 = .004$, BFinc = .14. There was also no significant interaction between content and perspective, $F(1, 39) < 1$, $p = .688$, $\eta_p^2 = .004$, BFinc = .06. Lastly, the effect of perspective was not significant, $F(1, 39) = 2.90$, $p = .096$, $\eta_p^2 = .069$, BFinc = .42.

The ANOVA for the left-handers showed a significant effect of content, $F(1, 23) = 4.59$, $p = .043$, $\eta_p^2 = .166$, BFinc = 4.64, indicating that the participants took longer to respond to the congruent items, compared to the incongruent items, which is in line with the accuracy scores for both left and right-handers, where it seemed that more correct answers were given on the incongruent items. There was no significant effect of perspective, $F(1, 23) < 1$, $p = .596$, $\eta_p^2 = .012$, BFinc = .19 nor a significant interaction between content and perspective, $F(1, 23) < 1$, $p = .752$, $\eta_p^2 = .004$, BFinc = .21. As the analysis for the left-handers showed no effect of perspective, only an analysis on participant's variability and not on item variability is reported here.

Table 1. Means (and standard deviations) of accuracy (in %) and reaction times (in msec).

	Condition			
	Content + perspective match	Content match + perspective mismatch	Content mismatch + perspective match	Content + perspective mismatch
Accuracy	94.3 (10.2)	94.0 (7.6)	99.2 (1.4)	99.2 (2.2)
Reaction time	498 (136)	519 (175)	488 (157)	505 (156)

ERP data

In Figure 2 the grand averaged ERPs for four electrodes (two frontal and two central electrodes) are shown. We chose to show these electrodes because, as mentioned earlier, some studies investigating sentence-picture integration have found N400-effects with a centro-frontal maximum (e.g., Federmeier & Kutas 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010).

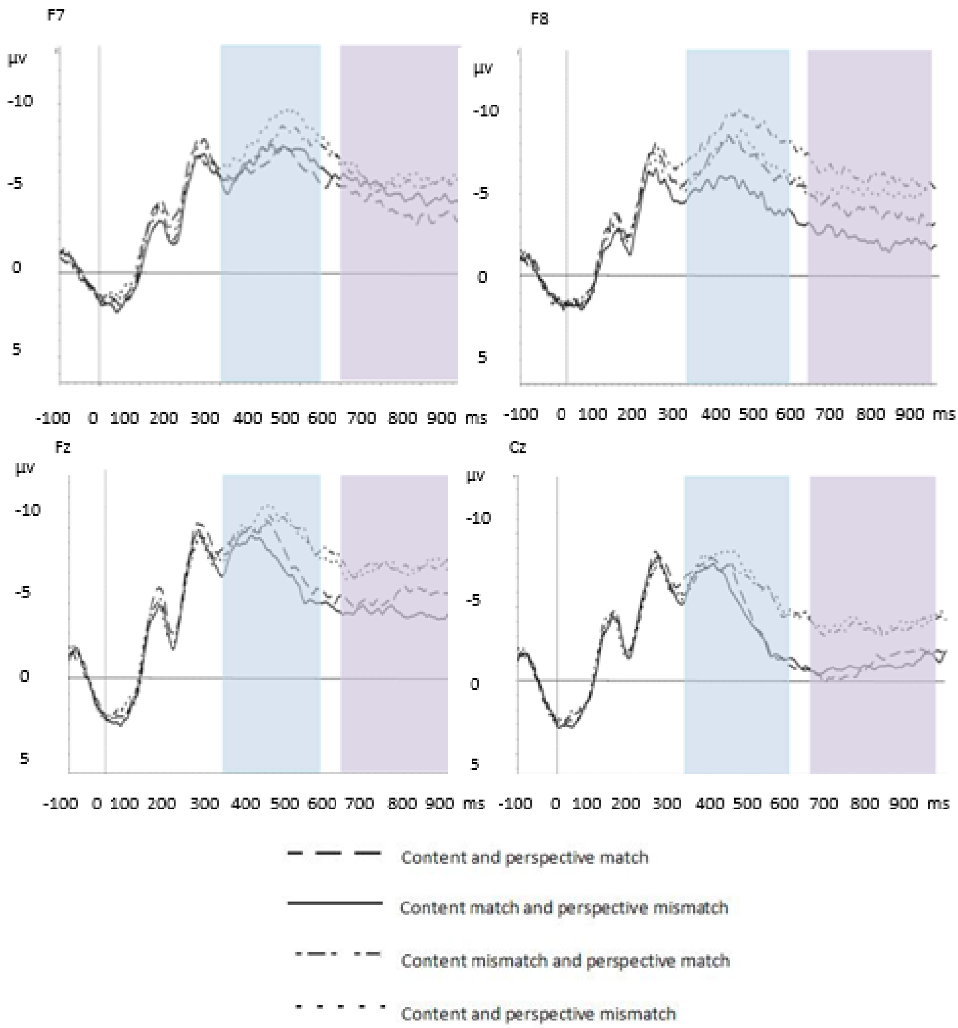


Figure 2. Grand –average ERPs for the four conditions at electrodes, F7, Fz, F8, Cz. Negativity is plotted upwards.

N400 time window (300-550 ms). In Figure 3, the grand-averaged ERPs are shown for the four conditions in the four quadrants used. The ANOVA showed a significant main effect of region, $F(4, 17) = 19.71$, $p < .001$, $\eta_p^2 = .82$, BFinc = ∞, meaning that there is a significant difference between the frontal regions and the posterior and midline regions ($p < .001$). This is, however, not relevant for our hypotheses, so we will not elaborate on these findings. There was also a significant effect of content, $F(1, 20) = 4.49$, $p = .047$, $\eta_p^2 = .18$, BFinc = 4.37, where we found a larger N400-amplitude for sentence-picture mismatching than for matching items. There was no effect of perspective, $F(1, 20) < 1$, $p = .951$, $\eta_p^2 < .001$, BFinc = .048. Given these data, the null hypothesis of no differences in performance on the perspective factor was about 21 times (inverse of the inclusion Bayes Factor) more likely than the alternative hypothesis, which could be described as strong evidence for the null hypothesis. Lastly, there was no interaction between content and perspective, $F(1, 20) = 1.49$, $p = .236$, $\eta_p^2 = .07$, BFinc = .057. Given these data, it is about 18 times more likely that the null hypothesis is true. It is, therefore, likely that there was truly no effect of perspective in the ERP data. Lastly, there was a significant interaction between region and perspective, $F(4, 17) = 4.28$, $p = .014$, $\eta_p^2 = .50$, BFinc = .005, although Bonferroni-corrected post-hoc tests showed no significant differences between the left and right-handed perspective picture conditions in the different regions (left anterior: $F(1, 20) = 1.06$, $p = .315$, BFinc = .33; right anterior: $F(1, 20) = 1.25$, $p = .28$, BFinc = .28; left posterior: $F(1, 20) < 1$, $p = .839$, BFinc = .22; right posterior: $F(1, 20) < 1$ $p = .893$, BFinc = .18; midline: $F(1, 20) < 1$, $p = .737$, BFinc= .26).

In the design of this study we used a certain amount of repetition of stimuli, which was necessary, given that there are only a limited number of manual action verbs. However, repetition of stimuli can influence the size of the N400 (Rugg, 1985), therefore we performed a Repeated-Measures ANOVA on only the first block of stimuli before any repetition occurred to test for the effect of repetition. Even though the power of this analysis is much lower than the overall analysis, it does not seem to be the case that there is an effect of perspective anywhere, $F(1, 19) < 1$, $p = .438$, $\eta_p^2 = .03$, BFinc = .13). Given these data, the null hypothesis of no differences on the perspective factor was about 8 times more likely than the alternative hypothesis, meaning there is substantial evidence for the null hypothesis.

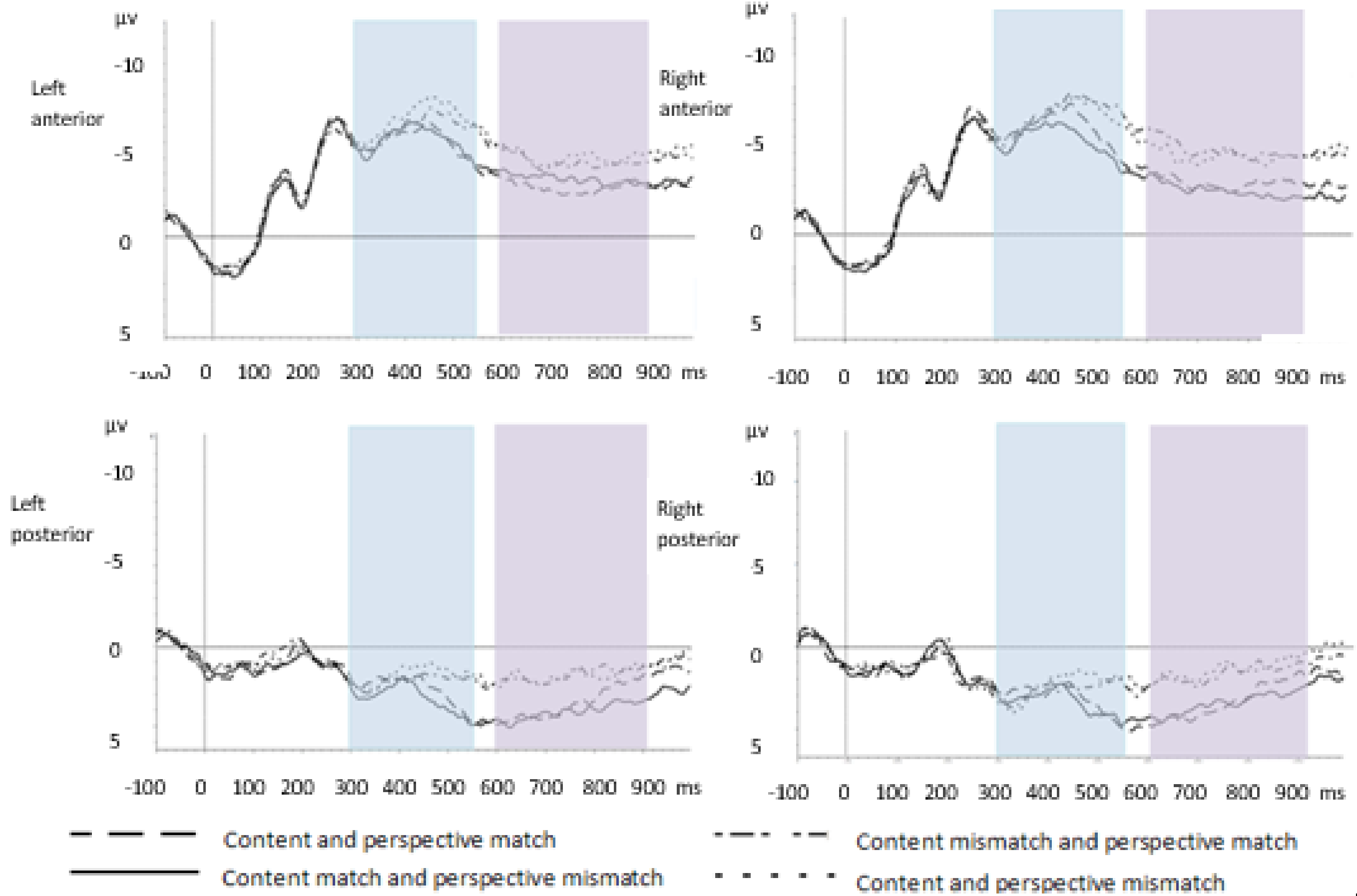


Figure 3. Grand-average ERPs for the four conditions for the four quadrants used. Negativity is plotted upwards.

Lastly, we conducted a Repeated-Measures ANOVA with hemisphere (left, right), content (congruent-incongruent) and perspective (left-right) as within-subject factors, as we were interested in any interactions of content or perspective with hemisphere. We found a significant interaction between hemisphere and perspective, $F(1, 20) = 6.14$, $p = .022$, $\eta_p^2 = .235$, BFinc = .02. Although it appeared there was a larger negativity for the left-handed pictures processed in the left hemisphere and for the right-handed pictures processed in the right hemisphere, than for the left-handed pictures processed in the right hemisphere and the right-handed pictures processed in the left hemisphere, both the Bayes factor (which suggests that the null hypothesis is 50 times more likely) and follow-up post hoc tests (left hemisphere: $t(1, 20) < 1$, $p = .546$, BFinc = .34; right hemisphere: $t(1 ,20) <1$, $p = .549$, BFinc = .35), suggest there is no effect of perspective in either the left or right hemisphere.

Late Time Window (600-900 ms). The ANOVA showed a main effect of region, $F(4, 17) = 36.20$, $p < .001$, $\eta_p^2 =.90$, BFinc = $3.217x10^{15}$, and a main effect of content, with a larger negative amplitude for the sentence-picture mismatching than for the matching items, $F(1, 20) = 13.72$, $p = .001$, $\eta_p^2 = .41$, BFinc = $3.269x10^{8}$. There was no effect of perspective, $F(1, 20) < 1$, $p = .991$, $\eta_p^2 < .001$, BFinc = .049, nor an interaction between content and perspective, $F(1, 20) < 1$, $p = .760$, $\eta_p^2 = .005$, BFinc = .035. Similar to the N400 time window, there was an interaction between region and perspective, $F(4, 17) = 5.14$, $p = .007$, $\eta_p^2 = .55$, BFinc = .005. As in the N400 time window, however, Bonferroni-corrected post-hoc tests showed no significant differences between the left and right-handed picture conditions in the different regions (left anterior: $F(1, 20) < 1$, $p = .393$, BFinc = .30; right anterior: $F(1, 20) < 1$, $p = .542$, BFinc = .23; left posterior: $F(1, 20) < 1$, $p = .752$, BFinc = .23; right posterior: $F(1,20) < 1$, $p = .763$, BFinc = .21; midline: $F(1, 20) < 1$, $p = .778$, BFinc = .22).

Lastly, like in the N400 time window the interaction between hemisphere and perspective was significant in the late time window. This analysis showed again a larger negativity for left-handed perspective pictures processed in the left hemisphere, and right-handed perspective pictures processed in the right hemisphere, than for right-handed pictures processed in the left hemisphere or left-handed perspective pictures processed in the right hemisphere. Like in the N400 time window, follow-up t-tests were not significant (left hemisphere: $t(1, 20) < 1$, $p = .667$, BFinc = .40; right hemisphere: $t(1 ,20) < 1$, $p = .805$, BFinc = .30). No other interactions with the factor hemisphere were significant.

Discussion

In this study we investigated the integration of semantic information conveyed through sentences and pictures and the influence of mental simulations herein. Our results show a larger N400-amplitude for the sentence-picture mismatching than for the matching items in the N400 time window. This effect was relatively broad in its scalp distribution, given that there was no interaction between the factors content and region. Although some studies using daily actions found a more frontally distributed bias (e.g., Federmeier & Kutas, 2001; Holcomb & McPherson, 1994; Mudrik et al., 2010; West & Holcomb, 2002), others do not (Friedrich & Friederici, 2004; Knoeferle et al., 2011), which seems to be more in line with our findings. Given that the N400-component occurs in response to violations of semantic expectancy, this study suggests that people attempt to integrate pictorial information presented concurrently with verbal information presented in a single sentence context, where they have more difficulty with integration when the sentence content does not match the picture content. This finding is in line with other studies in the field that have investigated integration of verbal and pictorial information (e.g., Ganis et al., 1996; Knoeferle et al., 2011; Nigam et al., 1992; Willems et al., 2008) and with the one-step model of language processing according to which every source of information, whether linguistic or extralinguistic can immediately constrain the interpretation of an utterance (e.g., Spivey-Knowlton, & Sedivy, 1995; Tanenhaus et al., 1995).

Concerning the results for the late time window (600-900 ms) we found a larger negativity when the content of the sentence and picture mismatched. Previous research showed an N400-effect when presenting participants with congruent or incongruent visual scenes, but also a more broadly pronounced distributed negativity which the authors attributed to late processes of semantic evaluation and response preparation (Mudrik et al., 2010). This would also explain our results, given that our participants were required to give a response on whether picture and sentence matched in content. In other studies on this topic, a response was not always required (e.g., Willems et al., 2008) and therefore this late effect might have been less strong in these studies.

So what can these results tell us about how verbal and visual information are integrated, what is the underlying mechanism? One possibility, which we mentioned before, is that pictures elicit a mental simulation, while the sentence itself also evokes a mental simulation. These perceptual representations might be integrated into a coherent message. The creation of two matching mental simulations could then facilitate the integration of pictures in a single sentence context. If mental simulations facilitate the integration of information from two modalities, we would however expect the integration to be hindered by pictures that mismatch the observer's mental simulation (in this case the observer's hand preference), but this was not evidenced by our results in the N400. Although we used stimuli containing the pronoun 'you' and first-person perspective pictures which both could have stimulated taking the actor's perspective in the simulation (Sato & Bergen, 2013), and even though the N400-effect seemed to be the strongest for the condition where the content and the hand perspective mismatched, there were no significant effects of hand perspective. Given that we did not instruct the participants on how to interpret the pronoun 'you', participants could have taken the perspective of a bystander, which might have affected their expectations about the handedness perspective. This seems unlikely, however, because it has been shown that the use of the second person pronoun leads to the adoption of the actor's perspective, without any instruction on how to interpret the pronoun (Sato & Bergen, 2013).

While the ERP results did not differ for perspective matches and mismatches, reaction times on the sentence-picture verification task did. This is in line with other behavioral tasks (i.e., learning and memory tasks), where an effect of these mental simulations was found (e.g., Apel et al., 2012; De Nooijer et al., 2013). Several studies have shown that right-handers' recall can have negative effects on memorization (e.g., Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009) and learning tasks (e.g., De Nooijer et al., 2013) when seeing the left-handed perspective, which may suggest that participants are influenced by the mismatching perspective in a conscious task, but not in automatic, fast, subconscious processes (as reflected in the N400). However, it could be the case that the faster reaction times of the right-handers on the right-handed perspective pictures are due to a type of motor priming, where responses with the right-hand were faster to right-handed perspective pictures. Another factor that could have played a role here is familiarity with the right-handed perspective, which might have caused faster processing of this perspective. Considering familiarity with the perspective, both left- and right-handers can be assumed to be more familiar with right-hander's actions as we most often observe other people's actions with a right-handed perspective (as 90% of the population is right-handed). However, with pictures from a first-person perspective, left-handers should be more familiar with the left-handed first-person perspective than with the right-handed perspective whereas they would be more familiar with the right-handed perspective when seeing pictures from the third-person perspective. If the results on the reaction time task are, therefore, due to motor priming or familiarity with the perspective, we would have expected left-handers (responding with their left hand) to react faster to the first-person left-handed perspective pictures. There was, however, no difference in reaction time between the left-handed and right-handed perspective pictures, in left-handers. However, given that we do not have data on what perspective left-handers are most familiar with, we cannot entirely rule out that familiarity had (some) effect in these results.

Lastly, could it be the case that mental simulations (in the form of a handedness mismatch) do have an influence in the integration process, but was undetected in this study? Could, for example, the lack of finding an effect of perspective be due to the implicit nature of the task concerning perspective? No explicit mention was given to the perspective manipulation while participants did have to judge whether the content of the picture matched the sentence, which was the other manipulation. It seems that mental simulations in essence occur outside of awareness. Any effect of mismatching mental simulation on the integration of information would, therefore, have been expected when no attention is drawn to the handedness perspective. Because of this reason it is unlikely that any handedness effects were overshadowed by the effects of congruency. To provide more evidence for this statement, we conducted a Bayesian analysis that suggested that, given our data, the null hypothesis is 21 times more likely than the alternative hypothesis, which provides strong evidence for that hypothesis (Jeffreys, 1961).

Effects of mental simulations on language processes, as described earlier, were found without explicit attention to the left-right difference. Therefore, if there was an effect of perspective, we would indeed especially have expected this to occur when no attention was drawn to the perspective manipulation. Another factor one might argue could have influenced the results is the use of a design in which the sentences and pictures were presented more than once. This repetition was necessary because we wanted to investigate one type of verb, namely object-manipulation verbs that elicit motor activation. Given the limited number of these types of verbs, a certain amount of repetition was necessary to obtain enough power. Although repetition can influence the size of the N400-effect (Rugg, 1985), this influence has been found in some tasks (e.g., lexical decision) but not in others (e.g., word/number discrimination) (Bentin & McCarthy, 1994). Also, although the sentences and pictures were repeated, they were never repeated in the same combination, for which reason the correspondence of the sentence and picture, still had to be evaluated for each stimulus, which, therefore differs from when the exact same word in isolation is repeatedly presented and the same evaluation has to be given by the participant (as was the case in for example, Rugg, 1990). Lastly, an analysis of the first block of the experiment, before any repetition occurred, did not give any indication that an effect of hand perspective might have gone unnoticed. We, therefore, think it is implausible that the null-effect of perspective is due to a type II error.

To summarize, this study suggests that photographs of manual actions can quickly be integrated into a single sentence context. Concerning the hand perspective of the pictures, there is no neurophysiological evidence that hand perspective seems to matter for the ease of integration into the sentence, although it might have effects on behavioral tasks, involving learning, memory or comprehension. This study contributes to the literature on how and when information from different modalities is integrated. Even in a single context verbal and visual information are easily integrated, but in this integration process mental simulations in the form of a mismatching hand perspective, might not play a role based on the neurophysiological evidence presented in this study or this role might be too small to be visualized in the N400.

▪ Chapter 9

Summary and general discussion

The studies presented in this dissertation focussed on how motoric information can influence learning and recalling of words. Learning novel words is an important part of acquiring a new language (e.g., Folse, 2014), whether it is a first, a second, or even a third language. It is, therefore, important to find effective instructional techniques that can improve the – sometimes cumbersome- process of learning new words. Of interest in this dissertation was the question of whether and if yes, how, activation of the motoric system during learning would be such a technique to support learning of novel words. The idea that motoric activation in the brain evoked by observing, imitating, or generating gestures might influence how words are learned, can be explained via the theory of embodied cognition.

This theory states that cognitive processes, like language and thought cannot be separated from how the body functions (e.g., Barsalou, 1999). For instance, when reading a word like 'writing' a mental simulation is created of this action, meaning that the motor experiences that the person has had with this action are reactivated. Proof for this idea comes from research showing that when people read a word like 'writing' this automatically leads to activation in the motor cortex of the brain. This means that the language system and the motor system can influence each other (e.g., Hauk, Johnsrude, & Pulvermüller, 2004). From an embodied view on word learning it would, therefore, be expected that activation of the motor system (in the form of mental simulations) leads to the creation of a richer representation of what has to be learned, because the information is presented in different modalities (both verbal and motoric). This might improve word learning by enriching the mental representation of the word (e.g., Goldin-Meadow, 2010). This embodied cognition perspective was the starting point for investigating to what extent the body can influence learning performance in this dissertation.

The goals of this dissertation were (1) to investigate what gesture instructions (observation, imitation, or enactment) are most effective for improving verb learning in primary school children, and whether these gestures are equally effective for learning verbs from different verb categories and for children with different levels of language ability? (2) to investigate whether the effectiveness of gestures in word learning is similar for children and adults. (3) to explore whether learning and memorizing action words is influenced by seeing depictions of actions that mismatch the mental simulation that is automatically created when hearing the definition of an action word. (4) try to explain where the potential effects of mental simulations on language processing stem from. The most important findings will be summarized in the following paragraphs.

Summary of the main findings

The study described in **Chapter 2** investigated what the best way of using gesturing is for children's learning of novel words in the first language. Previous research has suggested that learning can be boosted via gesture observation, observation and imitation, or enactment (i.e., self-generated actions; see e.g., Casale, 1985; Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011; Tellier, 2005). A direct comparison between these methods has never been made, although some studies suggest that active use of the body, via imitation or enactment, might be a stronger tool in word learning than merely observing a gesture (e.g., Tellier, 2007). In this study primary school children aged 8-9 years, learned novel abstract (e.g., to dismiss), locomotion (e.g., to stride), or object-manipulation verbs (e.g., to chisel). Children learned these words by hearing the verbal definition of the word twice. During the second presentation of the verbal definition they (1) did not see any gestures, (2) observed a gesture congruent with the meaning of the word (e.g., for the word 'to chisel' they saw a hitting motion towards a closed fist), (3) observed and imitated the gesture, (4) created and executed a gesture of their own.

Results showed that learning of locomotion verbs improved when observing congruent gestures compared to only hearing a verbal definition, but contrary to the hypothesis, learning of locomotion verbs did not improve when children had to imitate a gesture or when they created their own gesture. None of the gesture methods influenced learning of abstract verbs compared to only hearing a verbal definition. Interestingly, learning of object-manipulation verbs was only influenced by the use of gestures in children with high verbal abilities. In contrast, children with low verbal abilities seemed to be hindered by the use of gestures during learning. Children with low verbal abilities performed worse when using their own bodies (imitation or enactment) compared to when they learned the words via a verbal definition only. The effects described above were only present for performance scores on definition recall, but not on comprehension (measured via a fill-in-the gap test in which the word had to be filled in in a sentence context). In sum, the effectiveness of different gesture methods for word learning in primary school children seems to be dependent on verb type as well as language proficiency.

The study described in **Chapter 3** again investigated the influence of using gestures for word learning, focusing specifically on the effect of gesture imitation compared to observation only. In Chapter 2 the effects of gesture imitation over observation were not convincing or even absent, while there seem to be indications in the literature that gesture imitation improves test performance more than observation (Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011; Tellier, 2007; 2008). Therefore a direct comparison was made in this study between gesture observation and imitation in children one year older (nine- to ten-year-olds) than the children that participated in the study described in Chapter 2. In addition, a factor that might play a role in the effectiveness of gesture imitation over gesture observation in word learning was investigated, namely the moment at which a gesture is imitated. The research question was: What is the best moment to imitate a gesture? Is that during encoding, straight after hearing the definition of a novel word and having seen the congruent gesture (which could help storing the information in memory), or is this during retrieving, during test-taking, when the earlier observed gesture is retrieved from memory just before the definition of the words has to be written down on the test? Or maybe a combination of these methods leads to an increase in test performance (i.e., gesture imitation during both encoding and retrieval). In this case the context of learning and retrieval is the same, which might be effective in learning (i.e., Encoding Specificity Principle: Tulving & Osler, 1968).

Results showed that gesture imitation did not improve learning outcomes for abstract and locomotion verbs, compared to gesture observation. However, for object-manipulation verbs, gesture imitation led to better performance than observation when the gesture was imitated during either encoding or retrieval. When the gesture was imitated at both time points, gesture imitation was not more effective than gesture observation. At a delayed test after one week, only gesture imitation during retrieval led to higher performance scores than gesture observation. Again, there were no effects of gestures on a comprehension (i.e., fill-in-the-gap) test, but only on a recall test. In sum, imitating a gesture just before taking a test does not only facilitate retrieval of the earlier stored definition, but this successful retrieval can also lead to better recollection of the meaning of the new word after one week.

In the study reported in **Chapter 4**, the effects of gesture observation and imitation on word learning were again examined. In this study a comparison was made between the effect on learning in primary school children (between age nine and ten; note that overall, these children had lower language abilities than the children of the same age who participated in Chapter 3) and young adults, and given the findings from Chapter 2 and 3, only object-manipulation verbs were investigated. In addition, it was investigated whether the novelty of the object-manipulation verb of which the definition had to be learned had an influence on how effective the use of gesture imitation compared to gesture observation is for learning novel object-manipulation verbs. It was expected that gestures would support learning more when the verb (and the action it refers to) are novel to the learner because the gestures then provide additional (motoric) information, which can help with consolidation of the novel verb (e.g., 'hieuwen' means to pull up the anchor of a ship; note that there seems to be no direct translation for this verb available in English, in which one would say 'hauling in an anchor'). When learning the definition of an object-manipulation verb while both the word and the action are known (e.g., 'typen' [typing] means to form words by using a keyboard) and only the definition of the word needs to be learned, the gestures might become redundant (and it is known that presenting learners with redundant information does not help, and may even hinder their learning; Kalyuga & Sweller, 2014). Although it was expected that gesture imitation would lead to higher performance scores than gesture observation, in both children and adults, the effects of gesture imitation might be larger for children than for adults, as children's working memory capacity is still developing (Olesen, Macoveanu, Tegnér, & Klingberg, 2007). For this reason, the same task might be more difficult for children than for adults, meaning that the gestures could be expected to be more beneficial for learning the definition of both known and novel words for children than for adults. Research suggests that gestures are mostly beneficial in difficult tasks (McNeil, Alibali, & Evans, 2000).

Results did not show a main effect of gesture imitation. However, there was an interaction between age group (children vs. adults) and gesture imitation. It seemed that children scored equally well when observing or imitating a gesture during learning, while adults seemed to perform better when words (both novel and known) were learned in combination with gesture observation only. There was no interaction with word type (known, novel), suggesting that the effect of the type of gesture used during word learning was not influenced by prior knowledge of the word and action for which the definition had to be learned.

In the study reported in **Chapter 5** a different method was used to study the effect of motor activation on word learning. When hearing a description of a known action, a mental simulation is created of that action (as if executing the action oneself; Casasanto, 2009). In the three experiments described in this chaper, it was investigated whether seeing pictures of actions that are incongruent with the mental simulation that is made based on hearing the description of the meaning of the word (which is created automatically according to theories of embodied cognition) can influence word learning. In these studies adults learned a number of novel words from an artificial language in combination with pictures depicting an action from a right-handed, left-handed or bimanual first-person perspective. Earlier studies have shown that hearing a definition of an object-manipulation verb or seeing a picture of an action can lead to body-specific mental simulations (i.e., a right-hander creates a right-handed simulation: Casasanto, 2009). Therefore, seeing a picture of a left-hander performing an action with the left hand might create a conflict with a right-hander's mental simulation after hearing the verbal definition of a word. Because word meanings are in part made up of simulations of one's own actions (according to theories of embodied cognition), seeing a picture from a perspective incongruent with one's own mental simulation might, therefore, hinder the encoding and consolidation process of the new word.

Indeed, right-handers learned more verbal definitions, measured via a cued-recall task, when a word was coupled to a right-handed picture perspective or a bimanual perspective that matched the mental simulation, compared to seeing a left-handed perspective picture. For left-handers there was no difference in the number of words they learned via a left-handed, right-handed, or bimanual perspective picture. These results suggest that mismatches between pictures and mental simulations evoked by hearing action words can negatively affect right-handers' but not left-handers' word learning. Left-handers might have more experience with the right-handed perspective (as they live in a right-handed world) and might therefore not experience a mismatch when seeing a picture with a right-handed perspective.

In the study described in **Chapter 6** the boundary conditions of the effect from Chapter 5 were investigated, by studying whether it is dependent on observation perspective (first-person versus third-person perspective). Pictures from a third-person perspective have been found to be less embodied than pictures from a first-person perspective (Lorey et al., 2009). The first-person perspective is more tightly coupled to the sensory-motor system than the third-person perspective, which requires additional visuospatial processing (Jackson, Meltzoff, & Decety, 2006). The effects of handedness on right-handers' word learning might therefore be different in third-person perspective pictures than in first-person perspective pictures. In addition, it is investigated whether seeing left-handed perspective pictures in the first- and third-person perspective hinder learning novel words in right-handers, compared to learning without pictures.

Results of the first experiment showed that right-handers were not hindered by seeing a left-handed third-person perspective compared to a right-handed third-person perspective picture. The second experiment replicated the findings from Chapter 5 with regard to the first-person perspective (Exp. 2b) and the findings from Experiment 1 in Chapter 6 regarding the third-person perspective (Exp. 2a). Moreover, Experiments 2a and 2b showed that there was no difference in learning outcomes when words were learned with or without pictures, no matter which observation perspective (1^{st} or 3^{rd}) or handedness perspective (left or right) was depicted (i.e., seeing a right-handed perspective did not help learning and seeing a left-handed perspective picture did not hinder learning).

In the study reported in **Chapter 7** the effect of pictures conflicting with mental simulations was further investigated, by studying whether the negative effect of left-handed perspective pictures is also present when known verbs simply need to be remembered. In the first experiment, participants (young adults) had to remember 20 verbs from their first language. The object-manipulation verbs were presented on the screen for 5 seconds in combination with a picture in which the action that the verb referred to was depicted from a left-handed or right-handed first-person perspective. Participants either moved on to the next item immediately, or there was a two-second break in between the items, or participants were required to make a movement with their right hand to proceed to the next verb. Results showed that when there was a two-second break between the presentation of the items, right-handers remembered fewer words in a free recall task when they saw the picture with the left-handed compared to the right-handed perspective. When there was no break in between or when an arm movement was made, there was no difference in recall scores for the left or right-handed perspective pictures.

The results of Experiment 1 suggested that cognitive processes occurring during the two-second break in which the word and picture are no longer presented on screen, were responsible for the effect of combining a word with a left-handed or right-handed first-person perspective picture. Therefore, Experiment 2 investigated whether this effect might have been due to mental simulation during this two-second break, which would be reflected in looking behaviour in the absence of the stimuli (e.g., Johanson & Johanson, 2014). For this reason eye tracking was used in Experiment 2, to study initial processing of the stimuli and to find indications of rehearsal and mental imagery during the break in the absence of the stimuli that might explain the effect found in Experiment 1. Eye movement data showed that on relative dwell time (i.e., the sum of all gaze points within an area of interest divided by the total presentation time of the stimulus) during the 5 seconds the items were presented, there were no differences in looking behaviour between the left- and right-handed perspective pictures. First fixations seemed to be longer to the left-handed perspective pictures, but this difference was not statistically significant. During the inter-stimulus interval (i.e., the two-second break), however, looks to the spot where the items were previously presented were longer than to other areas. Initial looks to the spot where the item was previously presented were longer when the previous item was a picture from a right-handed perspective.

These results show that not only learning, but even recall of known action words can be affected by picture perspective, but only when a two-second break was present in between stimuli, during which the stimulus was not shown. The eye-tracking data suggest that this effect is due to mental imagery/retrieval practice during the break, as people looked back longer to the spot where the item was previously presented, which has been associated with retrieval processes (e.g., Altmann, 2004; Johansson & Johansson, 2014; Laeng & Teodorescu, 2000; Tremblay, Saint-Aubin, & Jalbert, 2006), when the item was paired with a right-handed picture. This suggests that rehearsal and mental imagery during the break in the absence of the stimuli, occurs less under a left-handed picture perspective, which may have interfered with memory consolidation.

In the study described in **Chapter 8** an EEG study was conducted to discover more about the nature of mental simulations in language processing. In this study participants heard sentences in which an action was described (e.g., you are slicing the tomato). In addition they were presented with pictures depicting a congruent action (slicing a tomato) or an incongruent action (ironing a shirt) with a right-handed or left-handed first-person perspective. In a sentence-picture verification task participants were asked to respond as soon as possible with their dominant hand to indicate whether the pictures matched the content of the sentence or not, while reaction times were measured. An EEG (electro encephalogram) was recorded during this task. The research question was whether seeing incongruent pictures that do not match the content of the sentence would lead to an N400-effect. This event-related potential (ERP) component reflects the neural mechanism of semantic integration in a (sentence)context, 400 ms. after the presentation of the crucial part of the stimulus (i.e., the verb in the sentence). When verbal and visual information cannot be integrated, a larger N400 peak is expected than when the information from these two sources can be integrated. It was studied whether seeing a picture that does not match the hand preference of the participant is also reflected in the N400. This study could therefore provide insights as to whether body-specific mental simulations play a role in the early process of integration of information from the verbal and visual modality.

Results showed that there was a significant N400-effect for content, meaning it is more difficult to integrate a picture into the context of the sentence when the picture does not match the content of the sentence. However, this effect was not modulated by handedness perspective. It is, therefore, unlikely that the integration of a picture into the sentence context is influenced by body-specific mental simulations. The results of the reaction time task, that was also conducted with left-handed participants at a later point in time, did show a handedness effect, however: right-handers were slower to react to left-handed pictures than to right-handed pictures, while left-handers reacted equally fast to left- and right-handed perspective pictures.

In the next sections the main results will be discussed and interpreted, and theoretical and practical implications of the results are given.

Part 1: The effects of gesture observation, imitation, and generation on word learning

What have we learned from the studies described in this dissertation about the effects of various uses of gestures (observation, imitation, generation) on children's word learning?

Effects differ according to verb type

First of all, the results suggest that the effects of gestures on learning performance are dependent on word type. In both Chapter 2 and 3, gestures did not improve learning of abstract verbs. This might be explained by looking at the motor activation that is required for executing the actions denoted by these words. The strength of the link to the motor system differs for abstract, locomotion, and object-manipulation verbs. While performance of the actions denoted by locomotion verbs requires leg activation, actions denoted by object-manipulation verbs require manual activation, and actions denoted by abstract verbs do not rely on overt motor activation for their execution. From an embodied view on word learning, it is expected that the use of gestures activates the motor system (in the form of mental simulations), which might improve learning by reducing working memory load, enriching the mental representation of the word, or both. As word meanings are also in part made up of motor information, using gestures that are congruent with what has to be learned could link the novel words to congruent motoric experiences, which could lead to deeper understanding of word meaning (i.e., richer encoding and better consolidation of the new word). However, if abstract words are not made up of motor information, then gestures might not contribute to the creation of a deeper understanding of the words, and therefore they might not aid in learning this type of word.

This explanation seems to fit with the results of both Chapter 2 and 3 as in those studies, none of the ways of using gesturing during learning improved learning of abstract words. Locomotion verbs, on the other hand, have a closer link to the motor system than abstract verbs, which would explain why, at least under a gesture observation condition in Chapter 2, learning outcomes increased for these verbs compared to hearing only a verbal definition. However, gesture imitation did not seem to improve learning of locomotion verbs. This result is in line with findings from Chapter 3, where locomotion verbs were not learned better under gesture imitation instructions compared to gesture observation instructions. Because action execution (i.e., action imitation and enactment) for these verbs consisted of getting up to perform the gestures, this might have broken concentration and focus on the verbal definition for these words, therefore not improving learning performance under these conditions.

In Chapter 3, it was found that only object-manipulation verbs were learned better when imitating a gesture compared to only observing a gesture. This could be explained by the strong link between language and manual gestures (Gentilucci & Corbalis, 2006), as both locomotion verbs and abstract verbs do not necessarily involve manual actions. In addition, abstract and locomotion verbs are less goal-directed than object manipulation verbs (words like 'to fear' or to 'to saunter' do not indicate a goal) and there is research that suggests that actions are hard to imitate when they are not goal-directed (Bekkering, Wohlschläger, & Gattis, 2000). Therefore, the effect of gesture imitation on learning of object-manipulation verbs can be stronger than for the other verb types. The results of Chapter 3 indicate that this is the case, but the findings from Chapter 2 and 4 suggest that this is only so when children have relatively high language abilities.

Effects differ according to children's language abilities

The effects of gesture imitation on learning object-manipulation verbs differed for children with high and low language abilities. Children with high language abilities (Chapter 2 and 3) seemed to benefit from gesture imitation during word learning, while children with low language abilities seemed to be hindered by imitating gestures (Chapter 2 and 4). Vocabulary knowledge is an important aspect of language ability and has been shown to be associated with working memory capacity (Daneman & Green, 1986). Children with higher language abilities presumably need less working memory capacity for the word learning task, leaving more capacity to imitate the gestures, which may explain why they benefitted from gesture imitation while children with lower verbal abilities did not. Simultaneously processing and imitating the observed object-manipulation gestures requires attentional and working memory resources on top of the resources required for processing and rehearsing the word meaning in relation to the action that is being observed and imitated. Children with higher language ability had sufficient cognitive capacity available to engage in imitation without hampering processing and rehearsal of the word meaning, while children with low language abilities did not. The added difficulty of simultaneously learning a word and thinking about how to imitate it could have been too challenging for the children with low verbal skills, especially, given that gestures for object-manipulation verbs are cognitively complex (Cartmill, Beilock, & Goldin-Meadow, 2012): the object (e.g., the chisel) needed to execute the action has to be imagined by the learner.

The children who participated in the study reported in Chapter 3 were approximately one year older than the children who participated in the study reported in Chapter 2. It can therefore be expected that overall, they would have higher language abilities than the children in Chapter 2 and would, therefore, need less of their working memory capacity to complete the word learning task, leaving more capacity to engage in imitating gestures. However, the children who participated in the study described in Chapter 4 were of the same age, but a large portion of them had low language abilities, which may explain why they did not benefit from gesture imitation during learning of object-manipulation words, as the children in this study had lower verbal abilities than the children (of the same age) described in Chapter 3.

Effects of imitation differ depending on the moment at which a gesture is imitated

Interestingly, Chapter 3 showed that when children have sufficiently high language abilities to benefit from imitation, the moment at which the gestures are imitated matters for the effect on test performance. Engaging in imitation after study improved immediate test performance, but imitation just prior to taking an immediate test, resulted not only in better immediate, but also in higher delayed test performance after one week. This means that gestures do not only affect encoding, but might also facilitate retrieval of items from the mental lexicon, which is in line with the Lexical Retrieval Hypothesis (Raucher et al., 1996) and with other studies that suggest that recall of verbal material improves when people are cued with self-generated hand gestures (e.g., Frick-Horbury, 2002). Concerning the type of gesture instruction that is most beneficial, it thus seems that for children gesture imitation is most beneficial, but only for learning object-manipulation words and only once the child has achieved a certain level of language proficiency. In this case the children need less of their working memory capacity to perform the word learning task, and have more capacity to engage in another task (imitating gestures).

Effects are measured on recall but not on fill-in-the gap comprehension tests

Another finding regarding the effectiveness of gesturing for word learning, that was consistent across the studies reported in Chapters 2 and 3, was that gestures did not affect comprehension test performance. This could mean that gesturing only affects recall; other studies investigating the effects of gestures on word learning mostly used free or cued-recall and recognition tests (e.g., Casale, 1985; Tellier, 2007) instead of comprehension tests. Another possibility, however, is that the fill-in-the-gap format of the comprehension test (although common in education) may not have been sensitive enough to measure effects of gesturing on comprehension. Other studies have shown that the *production* of new sentences with words that were learned, did improve when words were learned through enactment (although according to our definition, participants in this study used gesture imitation, as they imitated a gesture performed by an actor and they did not create a gesture themselves) compared to learning without gestures (Macedonia & Knösche, 2011), and this also requires comprehension. Moreover, comprehension of other learning processes seems to improve when gestures are used (for example in problem solving, understanding of novel concepts, math learning: Alibali, Spencer, Knox, & Kita, 2011; Broaders, Cook, Mitchell, & Goldin-Meadow 2007; Chu & Kita, 2011; Cook, Yip, & Goldin-Meadow 2012; Garber & Goldin-Meadow, 2002; Goldin-Meadow, Cook, & Mitchell 2009). So perhaps a production task might have provided us with a better measure of comprehension that could have resulted in different outcomes. This is something to address in future research. In sum, the studies described in this dissertation do not give an indication that the use of gestures influences performance on a comprehension task as measured by a fill-in-the-gap task.

Effects may differ for children and adults

The study reported in Chapter 4 suggested that adults did not benefit from gesture imitation over observation and that observation might even have been more effective. Because adults have sufficient cognitive capacity available for the task, it was expected that they would also benefit from gesture imitation, similar to the children with high language abilities, but this was not the case. At first sight, this seems to be in contrast to other studies in the field, which did find effects of gesture imitation over observation in second language learning in adults. In second language learning, action execution has been shown to enhance memory traces (e.g., Dijkstra & Kaschak, 2006; Engelkamp & Cohen, 1991) and improve word learning (e.g., Allen, 1995; Macedonia & Knösche, 2011; Macedonia, Müller, & Friederici, 2011). However, this was done in second language learning, whereas we had adults learn the definition of novel or known words from the first language (Chapter 4), which reflects a very different process. Adults might not benefit from gesture imitation over observation because adults have more years of perceptual and motor experience with words and actions than children and have more automated representations that come to mind with little effort, therefore they might not need to execute the gestures themselves to increase their word learning performance. Imitating the gestures might, thus, have been redundant for adults (and it is known that presenting learners with redundant information does not help, and may even hinder their learning; Kalyuga & Sweller, 2014). As no comparison was made with a verbal definition only condition, these data make it impossible to state whether gesture observation improves adults' word learning in the first language.

Perhaps, asking adults to enact the definitions they heard, without first observing a gesture, might have been more effective for them, as they would self-generate motor information and elaborate on the word meaning. Indeed, studies on the enactment effect did find memory advantages when participants acted out the action from the sentence (e.g., Engelkamp & Zimmer, 1997) and studies on multimedia learning often found benefits of generative learning activities such as drawing or explaining (Fiorella & Mayer, 2015). Future research might determine whether self-generated gestures would indeed be more effective for word learning in adults compared to gesture imitation; perhaps in relation to the effects of enactment on children's learning outcomes (which seemed absent in the study reported in Chapter 2).

Theoretical implications

In sum, the studies described in Chapters 2, 3, and 4 suggested that the effect of observing, imitating, and enacting gestures on verb learning are dependent on verb type, language proficiency, and age (with the latter two presumably being related to issues of working memory capacity). What do these findings on the use of gestures during word learning, mean for the development of theories on cognition? The results described in Chapter 2, 3, and 4 of this dissertation are not in line with the traditional view on cognition, which holds that cognition entails the manipulation of abstract, amodal symbols. Within this theory, perceptual and motoric systems are not necessary to understand cognitive processes. Instead, cognitive processing involves a small set of rules that is applied to a large set of amodal, abstract symbols (e.g., Fodor, 1983; Turing, 1950). The data presented in this dissertation, in contrast, do suggest that motoric information plays a role in cognitive processes like learning new words. The findings that observing or imitating gestures can improve word learning (Chapter 2, 3) would not be expected within this traditional view on cognition, and would also be hard to explain. On the other hand, the findings presented in this dissertation do not comply with a pure embodied cognition account either. In pure embodied accounts of cognition, it is claimed that cognition does not need amodal representations and that cognition is completely grounded in the sensorimotor system (e.g., Wilson, 2002). If this were the case, one would have expected that using gestures during learning would also improve learning of abstract words, but no such effects were found (Chapter 2, 3). The results are more in line with a less strict form of embodied cognition, wherein cognition is based on both sensori-motor and amodal information (e.g., Taylor & Zwaan, 2009; Zwaan, 2014; Zwaan, in press; Zwaan & Madden, 2005), and sensori-motoric information is thought to enrich amodal information and can lead to a deeper level of processing.

It should also be considered that there is another theory that can explain the findings reported here. Do the results provide evidence for embodied cognition, or can they also be explained by theories of multimedia learning (e.g., Mayer, 2014) and multisensory input (e.g., Shams & Seitz, 2008)? Both theories state that using congruent multimodal information (i.e., auditory / verbal, visual / pictorial, motoric) can lead to better learning than when presenting information in only one modality. When looking at the findings of the studies presented in Chapter 2 and 3, however, these are hard to explain from the multisensory theory or the cognitive theory of multimedia learning, because they show that adding motor information does not always benefit learning. Why should gestures have a differential effect on learning different word types? Not finding an effect for abstract verbs can be explained from an embodied cognition account, because of the weaker link these words have with the motor system. The effects of gestures on learning are therefore most likely not just due to adding an extra (motoric) modality during learning, but are more related to the sensori-motoric information that gestures offer, as the motoric information the gestures provide is congruent with what has to be learned.

Research implications

As mentioned above, children with low language abilities do not seem to benefit from imitating gestures during word learning, because they need all their working memory capacity for the word learning task, leaving no capacity to imitate the gestures and benefit from this extra information. One research implication could, therefore, be to practice the gestures that are combined with the novel words before the definitions are presented to these children. If the children are already familiar with the movements, executing them during a word learning task may require less working memory capacity. This could lead to children with low verbal abilities, being able to benefit from gesture imitation as well.

As it is important that children understand the meaning of the novel words they learn, future research could also focus on the use of different comprehension tests. The data from this dissertation did not give an indication that comprehension of the novel words improved when using gestures during word learning. Instead of a fill-in-the-gap test, other tests could be used to find out whether comprehension really does not improve when learning words in combination with gestures, or whether a fill-in-the-gap test is not sensitive enough to measure this effect.

Practical implications

For educational practice, these results qualify the suggestion that gesturing is effective for word learning in several important ways. First, whether any benefit can be expected from gestures depends on what type of verb the children have to learn, the kind of gesturing being used (observation only vs. imitation) and the children's language ability. Using gestures during abstract verb learning is probably not effective (although it does not seem to hamper learning either). Using gestures in verb learning seems most effective when learning concrete verbs, such as locomotion or object-manipulation verbs, although the type of gesturing that is effective seems to differ. When teaching children these concrete words, we have seen that gesture observation can best be used when teaching children locomotion verbs (e.g., to stride, to stroll), while gesture imitation seems to benefit learning of object-manipulation verbs (e.g., to mold, to chisel), provided that the children have a sufficiently high level of language ability.

Delaying the moment at which gestures are imitated can also have an effect on longer-term retention. Children were tested immediately after studying a block of five words, and when they imitated the gestures not immediately after each word, but just prior to the immediate test, their performance on a delayed test after one week was best. Although future research would have to investigate whether this also applies when the interval between study and test-taking is longer, this is something that teachers could use well when providing children with immediate practice tests during word learning, as a means to boost their longer-term retention.

Part 2: The effects of mental simulations on language processing and learning

The studies in the second half of this dissertation took a different approach to studying the effects of motor activation on language processing, by investigating whether visual information depicting an action that (mis)matches the learner's body-specific mental simulation, which we assume is automatically and unconsciously created when hearing (a definition of) a verb (e.g., Barsalou, 2008; Borghi et al., 2007; Chao & Martin, 2000; Grèzes & Decety, 2002) can influence word learning. The results reported in Chapters 5 and 6 seem to suggest that right-handers are indeed negatively influenced by learning words in combination with pictures that mismatch their body-specific mental simulation compared to matching pictures (i.e., they learned fewer definitions when these definitions were learned in combination with a left-handed compared to a right-handed perspective picture).

Effects of mental simulations during word learning and recall in right-handers

The small but consistent effect that right-handers learn or recall fewer verbal definitions of artificial words when these definitions are combined with left-handed than with right-handed first-person perspective pictures is in line with theories of embodied cognition, in which it is stated that both verbal information and pictures evoke mental simulations. When information from these two modalities is in contradiction, this can hinder the encoding and consolidation process of the new word. The results of the studies presented in Chapter 5 and 6 are also in line with the body-specificity hypothesis (Casasanto, 2009), which states that people create a body-specific mental simulation. Right-handers make right-handed embodied representations and are therefore differentially influenced by seeing a right- (matching) or left-handed (mismatching) picture perspective. The results are also in congruence with other studies that found a beneficial effect of seeing a picture of an object in an orientation that is congruent with a right-hander's mental simulation in memory tasks (e.g., Apel et al., 2012) and perceived distance tasks (Linkenauger et al., 2009). Also, when learning a knot-tying task, people learn the most when watching a model from a first-person perspective that has the same handedness (Michel & Harkins, 1985).

Although the explanation for all these findings lies most likely in (interference with) mental simulations, one possible alternative explanation is that the mismatching perspective leads to some kind of surprise effect, which could distract the learner and lead to lower performance scores. Future research could investigate whether this is a valid explanation or whether the effect can better be explained by (mis)matching mental simulations. For example, by comparing learning outcomes of left-handers when learning words for actions, for which they exclusively use their left hand, and for actions for which they use both their left and right hand. When the effects of Chapter 5 and 6 are due to (mis)matching body-specific mental simulations, it would be expected that left-handers learn words in combination with the left-handed perspective better for actions, for which they only use their left hand, as here they would also be expected to have strong hand-specific embodied connections. If the effects of Chapter 5 and 6 are merely due to a surprise effect, it is still hard to explain why there are differences in learning outcomes between right- and left-handers, but this might be investigated by comparing learning words in combination with right- and left-handed pictures and pictures that match or mismatch the content of the novel verb (e.g., seeing a staple when learning the word for chiselling). If people learn fewer words in combination with the left-handed picture because they are surprised to see this picture, then the same effect would be expected when seeing an unrelated picture. If the results differ for these types of pictures, then body-specific mental simulations might play a role in this learning process. In addition, it would be interesting to investigate what would happen when participants are simultaneously presented with the verbal definition and the picture. In this case the experienced mismatch might disappear, because there is no time to first create a simulation on the basis of the verbal definition, that then mismatches the mental simulation evoked by the picture. Here the two mental simulations might be unified. The idea that the effects found in Chapter 5 and 6 can be influenced by the circumstances in which the words and pictures are presented, is confirmed by the results from Chapter, 6, 7 and 8.

To illustrate, when seeing a third-person perspective during word learning there were no effects of handedness perspective, which was shown in two experiments described in Chapter 6. A reason for this could be that the first person-perspective is less embodied than the third-person perspective. This might mean that when a picture is viewed from a first-person perspective that this automatically leads to a motor imagery strategy, but while viewing the third-person perspective this might only lead to visual imagery (Sirigu & Duhamel, 2001), that would not necessarily affect the motor imagery evoked by the verbal definition. There is a possibility that seeing the third-person perspective might still influence learning, but only when the motor information is necessary to execute the task (Rohbanfard & Proteau, 2011).

In Chapter 7 two further conditions were revealed when the effect of mental simulations is not visible in memory: when there is no break between the items that have to be memorized and when an arm movement is made during the break between two items. Remembering words associated with a picture mismatching the mental simulation created by the observer can only hinder memory for those words in a free recall task, when there is time for such a simulation in absence of the stimuli (i.e., when there was a two-second break in between), a process that might be complemented by the conscious process of mental imagery. A period in which the information is retrieved before going on to the next item is a prerequisite for mismatching mental simulations to have an effect on memory. This is in line with recent suggestions that grounded congruency effects might be dependent on task conditions (Lebois, Wilson-Mendenhall, & Barsalou, 2014). However, why was there then no effect of mismatching mental simulations when an arm movement was made, also creating a break between the items? It could be that executing the arm movement to progress to the next item might have distracted participants' attention from rehearsing the item through imagery, and consequently, picture perspective might not have affected recall in this condition.

Lastly, body-specific mental simulations do not appear to play a role in the integration of a picture into a sentence context, as there was no effect of pictures that (mis)match the mental simulation on the N400 effect, measured via EEG, that reflects the neural mechanism of semantic integration into a context.

Effects of mental simulations during word learning in left-handers

As briefly discussed above, left-handers were not influenced by seeing the right-handed picture perspective during word learning (Chapter 5) nor when they had to respond as quickly as possible on a sentence-picture verification task where both left- and right-handed picture perspectives were used (Chapter 8). These findings are in line with other studies that found effects of handedness for right- but not left-handers. For example, left-handers did not recall and execute more instructions for moving objects when their handles were on the same side as their dominant hand (Apel et al., 2012) and left-handers do not perceive tools that were difficult to grasp as being farther away than easy to grasp tools, while right-handers did (Linkenauger et al., 2009). Left-handers were probably not influenced because they live in a right-handed world. They have more experience with seeing the right-handed perspective and they are sometimes forced to use their right hand as the world is organized to fit right-hander's needs. Left-handers might therefore not have very strong hand-specific embodied connections, as a consequence of which they do not experience a mismatch when seeing the right-handed perspective. This does not mean that left-handers are never influenced by their handedness. Several studies show, for example, that left-handers associate good things with the left side of space, while right-handers associate good things with the right side of space (e.g., Casasanto, 2009; Casasanto & Henetz, 2012; Scharine & McBeath, 2002). As mentioned above, an interesting question for future research would be whether left-handers can be influenced by seeing the right-handed picture perspective when learning words for actions for which they exclusively use their left-hand. In this case they would also be expected to have strong hand-specific embodied connections and might be influenced by seeing a perspective that mismatches their mental simulation.

Effects are measured on recall but not on comprehension tests

The effects of mental simulations on language processing and learning described above were only found on a cued-recall task (Chapter 5, 6) or free recall task (Chapter 7). However, on a translation task, (i.e., giving a translation in the first language of the artificial word, based on the verbal definition, which could indicate whether the participants really understood what the novel word means) used in the studies reported in Chapter 5 and 6, no effects were measured. This is in line with the studies reported in Chapter 2 and 3 where observing or imitating gestures did not affect performance on a comprehension test either. This could mean that that the effects of motoric information on language learning and processing are restricted to recall effects, or –as discussed above- it might be due to our particular choice of test. This is an issue that should be investigated in future research, with a different set of tests that could measure comprehension of the words that are learned.

Effects of learning with pictures compared to learning without pictures

Concerning the question whether pictures in general improve word learning, the results described in this dissertation are difficult to interpret, as seeing a right-handed perspective did not improve learning and seeing a left-handed picture perspective did not hinder learning compared to learning without pictures. Why is this? Research on the multimedia effect has provided some evidence that word learning can be improved when verbal information is combined with pictures (e.g., Chun & Plass, 1996; Jones, 2004; Kellogg & Howe, 1971; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968), although the findings from the studies reported in Chapter 6 are not in line with these results. The positive effect of providing pictures during word learning does not seem to be robust and effects appear to be larger for free recall tasks than for cued-recall tasks (e.g., Durso & Johnson, 1980; Lotto & De Groot, 1998), which were used in the study reported in Chapter 6 and could therefore have played a role in the lack of effect of pictures in word learning. In addition, it could be the case that the pictures in the studies described here were redundant as they show the same information that could be gained from the verbal definition; presenting redundant information is known to have no effect (as was the case here) or even a negative effect (which was not the case) on learning

(Kalyuga & Sweller, 2014). In addition, the task used in the study reported in Chapter 6 was completely verbal. Participants did not need the information presented in the pictures to do well on the test (as they were only asked to provide the verbal definition), again making the pictures redundant. If another task was used in which the information gleaned from the pictures was essential to perform the task, the pictures might have improved performance scores. This is a question for future research.

Theoretical implications

Again, the results of the second half of this dissertation are not in line with a pure embodied cognition view, where cognition involves only grounded symbols, results do favor a more moderate embodied cognition account over a traditional view of cognition in which only amodal symbols are used, which would make the results of Chapter 5 and 6 hard to explain. A more recent theory on cognition, the pluralist view, states we should not investigate whether we use abstract or grounded symbols for cognitive processes, but we should investigate when we use which symbols. The pluralist view suggests that cognitive processes use both grounded and abstract symbols (e.g., Dale, Dietrich, & Chemero, 2009; Dove, 2009; Zwaan, 2014), and that the question is not so much whether we make mental simulations, but under what circumstances. This can be dependent on the context, or the extent to which language comprehension is embedded in the environment. For example, when giving a demonstration (a cook tells you how to cut a bell pepper) perceptuo-motoric process will play an important role in understanding the cook’s message. In an instruction, like ‘give me the pencil’, a grounded representation of the pencil has to be activated (otherwise it would be hard to tell what object to look for). To understand an abstract conversation (e.g., a philosophical discussion), abstract symbols will be activated (Zwaan, 2014). Recently, it was argued that symbolic representations and the associations between them interact with sensorimotor representations to optimize discourse comprehension (Zwaan, in press).

Like the findings from the first part of this dissertation, the results from Chapters 5 and 6 are difficult to reconcile with a multisensory theory. Learning words by means of a verbal definition in which right-handed or left-handed people execute the action should have the same effect on learning, according to this theory, considering both convey information from a different modality. These findings can better be explained from the body-specificity hypothesis (Casasanto, 2009), discussed above.

In short, the studies discussed in this dissertation provided evidence for a view on cognition in which mental simulations are used, but where mental simulations are not always necessary for language comprehension and it depends on the context, when mental simulations influence cognitive processes, like learning novel words.

Research implications

The studies discussed above suggest that left-handers do not make body-specific mental simulations, or are not influenced by seeing a mental simulation that does not match how they execute an action, when learning or recalling words. Interesting for future research would be not only to see whether left-handers are influenced by seeing the right-handed perspective of actions for which they exclusively use their left-hand but also to investigate what happens when left-handers are taught completely novel actions. In this case they do not have visual or motoric experience with the novel action, and would only have created left-handed experiences with the action. It would then be expected, that seeing the novel action from the right-handed perspective, could influence their memory for this action.

Practical implications

The results presented above imply that the handedness of the learner should be taken into account when teaching adults novel words and recalling known words (Chapter 5, 6, 7). Even though 90% of the population is right-handed, there are some areas that can benefit from taking handedness into account (for a review on the implications of handedness in other areas, such as politics and marketing, see: De Nooijer & Willems, 2015). For example, when creating instructional videos, where a model is shown from a first-person perspective, it would be wise to choose a right-handed model, to optimise learning outcomes. In addition, when a right-handed teacher gives a demonstration (for instance how to tie your shoes), this can be done from both a first- and a third-person perspective, while a left-handed teacher could better choose to demonstrate the action from the third-person perspective.

Lastly, although using pictures in textbooks is a widely used method to improve learning, the study reported in this dissertation does not provide any evidence for the idea that using pictures during learning improves learning over a baseline in which no pictures are used. However, adding pictures during learning also does not hinder learning.

▪ References

Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology, 40,* 23-31. doi: 10.1111/j.1467-8535.2007.00800.x

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science, 21,* 1903-1909. doi: 10.1177/0956797610389192

Alaerts, K., Heremans, E., Swinnen, S. P., & Wenderoth, N. (2009). How are observed actions mapped to the observer's motor system? Influence of posture and perspective. *Neuropsychologia 47,* 415–422. doi: 10.1016/j.neuropsychologia.2008.09.012

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science, 22*, 1138-1144. doi:10.1177/09567 97611417722

Allen, L. Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal, 79,* 521-529. doi: 10.1111/j.1540-4781.1995.tb05454.x

Altmann, G. (2004). Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm'. *Cognition, 93,* B79-B87. doi: 10.1016/j.cognition.2004.02.005

Amoruso, L., Gelormini, C., Aboitiz, F., González, M. A., Manes, F., Cardona, J. F., & Ibanez, A. (2013). N400 ERPs for actions: building meaning in context. *Frontiers in Human Neuroscience, 7*: 57. doi: 10.3389/fnhum.2013.00057

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology, 103,* 62-70. doi: 10.1037/h0036799

Apel, J. K., Cangelosi, A., Ellis, R., Goslin, J., & Fischer, M. H. (2012). Object affordance influences instruction span. *Experimental Brain Research, 223,* 199-206. doi: 10.1007/s00221-012-3251-0

Asher, J. J., & Price, B. S. (1967). The learning strategy of the total physical response: Some age differences. *Child Development, 38,* 1219-1227.

Atkinson, R. C. (1975). Mnemotechinics in second-language learning. *American Psychologist, 30,* 821-828. doi:10.1037/h0077029.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70,* 181-214. doi: 10.3102/00346543070002181

Avikainen, S., Kulomäki, T., & Hari, R. (1999). Normal movement reading in Asperger subjects. *Neuroreport, 10,* 3467-3470. doi: 10.1097/00001756-199911260-00001

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology, 16,* 1818-1823. doi: 10.1016/j.cub.2006.07.060

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database*.

Bach, P., Gunter, T. C., Knoblich, G., Prinz, W., & Friederici, A. D. (2009). N400-like negativities in action perception reflect the activation of two components of an action representation. *Social Neuroscience, 4,* 212-232. doi: 10.1080/17470910802362546

Baddeley, A. D. (1999). *Essentials of human memory*. Psychology Press.

Bandura, A. (1986*). Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22,* 577-660. doi: 10.1017/S0140525X99532147

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

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58*, 495-520. doi: 10.1016/j.jml.2007.02.004

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences, 105,* 13269-13273. doi: 10.1073/pnas.0803424105

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology, 53A*, 153–164. doi: 10.1080/713755872

Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 130-149. doi: 10.1037/0278-7393.20.1.130

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication. *Neuropsychologia, 44,* 178–190. doi: 10.1016/j.neuropsychologia.2005.05.007

Borghi, A. M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., & Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information? Studies with hand primes. *Neuroscience Letters, 411,* 17-21. doi: 10.1016/j.neulet.2006.10.003

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory and Cognition, 32,* 863-873. doi: 10.3758/BF03196865

Borghi, A. M. (2005). Object concepts and action. In D. Pecher & R. A. Zwaan (Eds.), *The grounding of cognition: The role of perception and action in memory, language, and thinking* (pp. 8-34). Cambridge: Cambridge University Press. doi : 10.1017/CBO9780511499968.002

Borghi, A. M., Caramelli, N., & Setti, A. (2005). Conceptual information on objects' locations. *Brain and Language, 93,* 140-151. doi: 10.1016/j.bandl.2004.09.004

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science, 13,* 185-189. doi: 10.1111/1467-9280.00434

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9,* 27-38. doi: 10.1162/jocn.1997.9.1.27

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136,* 539. doi: 10.1037/0096-3445.136.4.539

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5,* 34-44. doi: 10.1162/jocn.1993.5.1.34

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high quality, data? *Perspectives on Psychological Science, 6,* 3–5. doi: 10.1177/1745691610393980

Busiello, M., Costantini, M., Galati, G., & Committeri, G. (2011). Sensory-motor interference abolishes repetition priming for observed actions, but not for action-related verbs. *Neuroscience Letters, 492,* 89-93. doi: 10.1016/j.neulet.2011.01.063

Cartmill, E. A., Beilock, S., & Goldin-Meadow, S. (2012). A word in the hand: action, gesture and mental representation in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 129-143. doi: 10.1098/rstb.2011.0162

Casale, U. P. (1985). Motor imaging: A reading-vocabulary strategy. *Journal of Reading, 28,* 619-621.

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

Casasanto, D., & Henetz, T. (2012). Handedness shapes children's abstract concepts. *Cognitive Science, 36,* 359-372. doi: 10.1111/j.1551-6709.2011.01199.x

Chao, L. L., & Martin, A. (2000) Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12,* 478–84. doi: 10.1006/nimg.2000.0635

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General, 137*, 706. doi: 10.1037/a0013157

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General, 140,* 102. doi: 10.1037/a0021790

Chun, D., & Plass, J. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal, 80,* 183-198. doi: 10.1111/j.1540-4781.1996.tb01159.x

Cito (2009). *Monitor taal en rekenen. Eerste meting: een indicatie van leerprestaties in termen van het referentiekader* [Monitor language and calculus. First measurement: an indication of learning abilities in terms of the frame of reference]. Arnhem: Cito.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised edition). New York: Academic Press.

Cohen, M. T., & Johnson, H. L. (2011). Improving the acquisition of novel vocabulary through the use of imagery interventions. *Early Childhood Education Journal, 38,* 357-366. doi: 10.1007/s10643-010-0408-y.

Cohen, R. L., & Otterbein, N. (1992). The mnemonic effect of speech gestures: Pantomimic and non-pantomimic gestures compared. *European Journal of Cognitive Psychology, 4,* 113-139. doi: 10.1080/09541449208406246

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497-505. doi: 10.1080/14640748108400805

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6,* 256-266. doi: 10.1162/jocn.1994.6.3.256

Cook, S. W., & Fenn, K. M. (2010). *Gesture improves retention and transfer of symbolic learning in adults.* Paper presented at the fourth International Society for Gesture Studies, Frankfurt an der Oder, Germany. Abstract retrieved from http://www.gesturestudies.com/ISGS10 Abstracts.pdf

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106,* 1047-1058. doi: 10.1016/j.cognition.2007.04.010

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes, 27,* 594-610. doi: 10.1080/01690965.2011.567074

Coren, S., & Porac, C. (1977). Fifty centuries of right-handedness: The historical record. *Science, 198,* 631–632. doi: 10.1126/science.335510

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268–294. doi:10.1037/0096-3445.104.3.268

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* Routledge.

D’Arcy, R. C. N., & Connolly, J. F. (1999). An event-related brain potential study of receptive speech comprehension using a modified Token Test. *Neuropsychologia, 37,* 1477-1489. doi: 10.1016/S0028-3932(99)00057-3

Dale, R., Dietrich, E., & Chemero, A. (2009). Explanatory pluralism in cognitive science. *Cognitive Science, 33,* 739-742. doi: 10.1111/j.1551-6709.2009.01042.x

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1-18. doi: 10.1016/0749-596X(86)90018-5

Decety, J., & Jeannerod, M. (1995). Mentally simulated movements in virtual reality: does Fitt's law hold in motor imagery? *Behavioural Brain Research, 72,* 127-134. doi: 10.1016/0166-4328(96)00141-6

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica*, 1*44*, 173-179. doi: 10.1016/j.actpsy.2013.05.013

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives. *Psychological Science, 24,* 2515-2521. doi: 10.1016/j.actpsy.2013.05.013

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2014). Words in action: Using gestures to improve verb learning in primary school children. *Gesture, 14,* 47–70. doi 10.1075/gest.14. 1.03noo

De Nooijer, J. A., & Willems, R. M. (in press). What can we learn about cognition from studying handedness? Insights from cognitive neuroscience. In N. Hagemann, B. Strauss, C. MacMahon & F. Loffing (Eds.). *Laterality in Sports: Theories and Applications.* San Diego, California: Academic Press, Elsevier.

De Ruiter, J. -P. (1998). *Gesture and speech production.* (Doctoral dissertation). Max Planck Institute for Psycholinguistics.

Dijkstra, K., & Kaschak, M. P. (2006). Encoding in verbal, enacted and autobiographical tasks in young and older adults. *Quarterly Journal of Experimental Psychology, 59,* 1338-1345. doi: 10.1080/17470210600695092

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition, 102,* 139-149. doi: 10.1016/j.cognition.2005.12.009

Dove, G. O. (2009). Beyond perceptual symbols: a call for representational pluralism. *Cognition, 110,* 412–431. doi: 10.1016/j.cognition.2008.11.016

Durso, F. T., & Johnson, M. K. (1980). The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning and Verbal Behavior, 19,* 416-429. doi: 10.1016/S0022-5371(80)90294-7

Eerland, A., Guadalupe, T. M., & Zwaan, R. A. (2011). Leaning to the left makes the eiffel tower seem smaller: posture-modulated estimation. *Psychological Science, 22,* 1511-1514. doi: 10.1177/0956797611420731

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42*, 1259-1268. doi: 10.1037/0012-1649.42.6.1259

Elder, R. S., & Krishna, A. (2012). The 'Visual Depiction Effect' in advertising: Facilitating embodied mental simulation through product orientation. *Journal of Consumer Research, 38,* 988-1003. doi: 10.1086/661531

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology, 49A,* 234-250. doi: 10.1080/713755604

Engelkamp, J. (1998). *Memory for actions.* Psychology Press, Hove, UK.

Engelkamp, J., & Cohen, R. L. (1991). Current issues in memory of action events. *Psychological Research, 53,* 175-182. doi: 10.1007/BF00941384

Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory unit. *Psychological Research, 46*, 283-299. doi: 10.1007/BF00308889

Engelkamp, J., & Zimmer, H. D.- (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica, 96,* 43-60. doi: 10.1016/S0001-6918(97)00005-X

Engelkamp, J., Zimmer, H. D., Mohr, G., & Sellen, O. (1994). Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition, 22,* 34-39. doi: 10.3758/BF03202759

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1-11. doi: 10.3758/BF03203630

Ernest, C. H. (1977). Imagery ability and cognition: A critical review. *Journal of Mental Imagery, 1,* 181-215.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 202-224. doi: 10.1037/0278-7393.27.1.202

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences, 12*, 405-410. doi: 10.1016/j.tics.2008.07.007

Feyereisen, P. (2009). Enactment effects and integration processes in younger and older adults' memory for actions. *Memory, 17,* 374-385. doi: 10.1080/09658210902731851

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Fiorella, L., & Mayer, R. E. (2014). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of motor system in language comprehension. *Quarterly Journal of Experimental Psychology, 61,* 825-850. doi: 10.1080/17470210701623605

Folse, K. S. (2004). *Vocabulary Myths: Applying second language research to classroom teaching.* Ann Arbor: University of Michigan Press.

Frick-Horbury, D. (2002). The effects of hand gestures on verbal recall as a function of high-and low-verbal-skill levels. *The Journal of General Psychology, 129*, 137-147. doi: 10.1080/00221 300209603134

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience, 16,* 1465-1477. doi: 10.1162/0898929042304705

Gallese, V., Gernsbacher, M. A., Heyes, C., Hickok, G., & Iacoboni, M. (2011). Mirror neuron forum. *Perspectives on Psychological Science, 6,* 369-407. doi: 10.1177/1745691611413392.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22,* 455-479. doi: 0.1080/026432904420 00310

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8,* 89-106. doi: 10.1162/jocn.1996.8.2.89

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science, 26*, 817-831. doi: 10.1016/S0364-0213(02)00087-3

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews, 30,* 949–960. doi: 10.1016/j.neubiorev.2006.02.004

Gibson, J. J. (1977). *The theory of affordances*. Hilldale, USA.

Glenberg, A. M. (1997). What memory is for. *Behavioral & Brain Sciences, 20,* 1–55. doi: 10.1017/S0140525X97000010

Glenberg, A. M. (2011). Introduction to the mirror neuron forum. *Perspectives on Psychological Science, 6*, 363-368. doi: 10.1177/1745691611412386

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

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61,* 905-919. doi: 10.1080/17470210701625550

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325–331. doi: 10.1111/j.2044-8295.1975.tb01468.x

Goodwyn, S. W., Acredolo, L. P., & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior, 24,* 81-103. doi: 10.1023/A:1006653828895

Goldin-Meadow, S. (1999). The role of gestures in communication and thinking. *Trends in Cognitive Sciences, 3,* 419–429. doi: 10.1016/S1364-6613(99)01397-2

Goldin-Meadow, S. (2010). When gesture does and does not promote learning. *Language and Cognition*, 2, 1-19. doi: 10.1515/langcog.2010.001.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20,* 267-272. doi: 10.1111/j.1467-9280.2009.02297.x

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*, 516-522. doi: 10.1111/1467-9280.00395

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neuropsychology, 55,* 468–484. doi: 10.1016/0013-4694(83)90135-9

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia 40,* 212–222. doi: 10.1016/S0028-3932(01)00089-6

Grèzes, J., Frith, C., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage, 21,* 744-50. doi: 10.1016/S1053-8119(03)00665-7

Guo, K., Mahmoodi, S., Robertson, R. G., & Young, M. P. (2006). Longer fixation duration while viewing face images. *Experimental Brain Research, 171,* 91-98. doi: 10.1007/s00221-005-0248-y

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438-441. doi: 10.1126/science.1095455

Hald, L. A., De Nooijer, J. A., Van Gog, T., & Bekkering, H. (2015). *Optimizing word learning via links to perceptual and motoric experience.* Manuscript submitted for publication.

Hald, L., Marshall, J., Janssen, D. P., & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology: Cognition, Special issue on Embodied and Grounded Cognition, 2*: 45. doi: 10.3389/fpsyg.2011.00045

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology, 113,* 1339-1350. doi: 10.1016/S1388-2457(02)00161-X

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

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science, 21,* 895-900. doi: 10.1177/0956797610374742

Hesse, M. D., Sparing, R., & Fink, G. R. (2009). End or means—the ‘‘What’’ and ‘‘How’’ of observed intentional actions. *Journal of Cognitive Neuroscience, 21,* 776–790. doi: 10.1162/jocn.2009.21058

Hickok, G. (2010). The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes, 25,* 749-776. doi: 10.1080/01690961003595572

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta analysis. *Learning and Instruction, 17*, 722-738. doi: 10.1016/j.learninstruc.2007.09.013

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition, 24,* 259–276. doi: 10.1006/brcg.1994.1014

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford University Press.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137,* 297. doi: 10.1037/a0022128

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396,* 228-228. doi: 10.1038/24300

Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language, 28*, 164-181. doi: 10.1177/0142723707087736

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage, 31,* 429-439. doi: 10.1016/j.neuroimage.2005.11.026

James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensorimotor systems in the developing brain. *Developmental Science 14,* 673–678. doi: 10.1111/j.1467–7687.2010.01011.x

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences, 17,* 187-201. doi: 10.1017/S0140525X00034166

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia, 33*, 1419-1432. doi: 10.1016/0028-3932(95)00073-C

Jeffreys, H. (1961). *The theory of probability*. Oxford: The Clarendon Press.

Jennings, J. R., Cohen, M. J., Ruchkin, D. S., & Fridlund, A. J. (1987). Editorial policy on analyses of variance with repeated measures. *Psychophysiology, 24,* 474–478.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science, 25,* 236-242. doi: 10.1177/0956797613498260

Johansson, R., Holsanova, J., & Holmqvist, K. (2005). *What do eye movements reveal about mental imagery? Evidence from visual and verbal elicitations*. In Proceedings of the 27th Cognitive Science Conference (p. 1054).

Jones, L. C. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning and Technology, 8,* 122-143.

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (pp. 247-262). New York: Cambridge University Press.

Kellogg, G. S., & Howe, M. J. (1971). Using words and pictures in foreign language learning. Alberta *Journal of Educational Research, 17,* 89-94.

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40,* 577-592. doi: 10.1006/jmla.1999.2634

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience, 22,* 683-694. doi: 10.1162/jocn.2009.21254

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language, 89,* 253-260. doi: 10.1016/S0093-934X(03)00335-3

Kent, C., & Lamberts, K. (2008). The encoding–retrieval relationship: retrieval as mental simulation. *Trends in Cognitive Sciences, 12*, 92-98. doi: 10.1016/j.tics.2007.12.004

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*, 313-334. doi: 10.1080/01690960802365567

Keysers, C., & Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology, 19,* 666-671. doi: 10.1016/j.conb.2009.10.006

Kiefer, M., & Trumpp, N. M. (2012). Embodiment theory and education: the foundations of cognition in perception and action. *Trends in Neuroscience and Education, 1,* 15-20. doi: 10.1016/j.tine.2012.07.002

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.)*, Language and gesture* (pp. 162–185). Cambridge, UK: Cambridge University Press.

Klöppel, S., Van Eimeren, T., Glauche, V., Vongerichten, A., Münchau, A., Frackowiak, R. S., Büchel, C., Weiller, C., & Siebner, H. R. (2007). The effect of handedness on cortical motor activation during simple bilateral movements. *Neuroimage, 34,* 274-280. doi: 10.1016/j.neuroimage. 2006.08.038

Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. *Psychophysiology, 48,* 495-506. doi: 10.1111/j.1469-8986.2010.01080.x

Kormi-Nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology, 7,* 337-363. doi: 10.1080/0954144 9508403103

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205. doi: 10.1126/science.7350657

Laeng, B., & Teodorescu, D-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26,* 207–231. doi: 10.1016/S0364-0213(01)00065-9

Lebois, L. A. M., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2014). Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science,* 1-38. doi: 10.1111/cogs.12174

Lewis, J. W., Phinney, R. E., Brefczynski-Lewis, J. A., & DeYoe, E. A. (2006). Lefties get it “right” when hearing tool sounds. *Journal of cognitive neuroscience, 18,* 1314-1330. doi: 10.1162/jocn.2006.18.8.1314

Linkenauger, S. A., Witt, J. K., Stefanucci, J. K., Bakdash, J. Z., & Proffitt, D. R. (2009). The effects of handedness and reachability on perceived distance. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1649-1660. doi: 10.1037/a0016875

Liuzza, M. T., Setti, A., & Borghi, A. M. (2012). Kids observing other kids' hands: Visuomotor priming in children. *Consciousness and Cognition, 21,* 383–392. doi: 10.1016/j.concog.2011.09.015

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 565-572. doi: 10.1037/0096-1523.4.4.565

Longcamp, M., Anton, J. L., Roth, M., & Velay, J. L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage, 19,* 1492-1500. doi: 10.1016/S1053-8119(03)00088-0

Lorey, B., Bischoff, M., Pilgramm, S., Stark, R., Munzert, J., & Zentgraf, K. (2009). The embodied nature of motor imagery: the influence of posture and perspective. *Experimental Brain Research, 194,* 233-243. doi: 10.1007/s00221-008-1693-1

Lotto, L., & De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in a new language. *Language Learning, 48,* 31-69. doi: 10.1111/1467-9922.00032

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 227-246). New York: Cambridge University Press

Macedonia, M. (2003). *Sensorimotor enhancing of verbal memory through "Voice Movement Icons" during encoding of foreign language* (Doctoral dissertation). Salzburg: University of Salzburg.

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education, 5,* 196-211. doi: 10.1111/j.1751-228X.2011.01129.x

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping, 32,* 982-998. doi: 10.1002/hbm.21084

Madan, C. R., & Singhal, A. (2012). Encoding the world around us: Motor-related processing influences verbal memory. *Consciousness and Cognition, 21,* 1563–1570. doi: 10.1016/j.concog.2012.07.006

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. doi: 10.1016/j.jphysparis.2008.03.004

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64,* 17-24. doi: 10.1111/j.2044-8295.1973.tb01322.x

Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research, 77,* 303-309. doi: 10.1007/s00426-012-0439-7

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57,* 203–220. doi: 10.1037/h0087426

Mayer, R. E. (Ed.) (2014). *The Cambridge handbook of multimedia learning.* New York: Cambridge University Press.

McGregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). Gesture as a support for word learning: *The case of under. Journal of Child Language, 36,* 807-828. doi: 10.1017/S0305000908009173

McGregor, K. K., Sheng, L., & Ball, T. (2007). Complexities of expressive word learning over time. *Language, Speech, and Hearing Services in Schools, 38,* 353-364. doi: 10.1044/0161-1461(2007/037)

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior, 24*, 131-150. doi: 10.1023/A:1006657929803

McNeill, D, & Levy, E. (1982). Conceptual representations in language activity and gesture. In R .J. Jarvella, & W. Klein (Eds.), *Speech, place and action*. John Wiley and Sons, New York.

Mecklinger, A., Gruenewald, C., Weiskopf, N., & Doeller, C. F. (2004). Motor affordance and its role for visual working memory: evidence from fMRI studies. *Experimental Psychology, 51,* 258 –269. doi: 10.1027/1618-3169.51.4.258

Michel, G. F., & Harkins, D. A. (1985). Concordance of handedness between teacher and student facilitates learning manual skills. *Journal of Human Evolution, 14,* 597–601. doi: 10.1016/S0047-2484(85)80085-3

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156. doi: 10.1037/0022-0663.94.1.156

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology, 117,* 411-424.

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia, 48,* 507-517. doi: 10.1016/ j.neuropsychologia.2009.10.011

Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18,* 1098-1111. doi: 10.1162/ jocn.2006.18.7.1098

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4,* 15-22. doi: 10.1162/jocn.1992.4.1.15

Nilsson, L-G. (2000). Remembering actions and words. In E. Tulving, & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 137-148). New York, NY, US: Oxford University Press.

Nyström, P., Ljunghammar, T., Rosander, K., & Von Hofsten, C. (2011). Using mu-rhythm desynchronization to measure mirror neuron activity in infants. *Developmental Science, 14,* 327-335. doi: 10.1111/j.1467-7687.2010.00979.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9,* 97-113. doi: 10.1016/0028-3932(71)90067-4

Olesen, P. J., Macoveanu, J., Tegnér, J., & Klingberg, T. (2007). Brain activity related to working memory and distraction in children and adults. *Cerebral Cortex, 17,* 1047-1054. doi: 10.1093/cercor/ bhl014.

Olsen, R. K., Chiew, M., Buchsbaum, B. R., & Ryan, J. D. (2014). The relationship between delay period eye movements and visuospatial memory. *Journal of Vision, 14,* 8.doi: 10.1167/14.1.8

Osman, M., Bird, G., & Heyes, C. (2005). Action observation supports effector-dependent learning of finger movement sequences. *Experimental Brain Research, 165,* 19-27. doi: 10.1007/s00221-005-2275-0

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5,* 176–206. doi: 10.1016/0010-0285(73)90032-7

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11,* 137–138. doi: 10.3758/BF03331011

Paulus, M., Lindemann, O., & Bekkering, H. (2009). Motor simulation in verbal knowledge acquisition. *Quarterly Journal of Experimental Psychology, 62,* 2298-2305. doi: 10.1080/1747021090 3108405

Pecher, D. (2013). No role for motor affordances in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 2–13. doi: 10.1037/a0028642

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

Pecher, D., De Klerk, R. M., Klever, L., Post, S., Van Reenen, J. G., & Vonk, M. (2013). The role of affordances for working memory for objects. *Journal of Cognitive Psychology, 25,* 107-118. doi: 10.1080/20445911.2012.750324

Ping, R. M. & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44,* 1277-1287. doi: 10.1037/0012-1649.44.5.1277

Post, L. S., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of simultaneously observing and making gestures while studying grammar animations on cognitive load and learning. *Computers in Human Behavior, 29,* 1450-1455. doi: 10.1016/j.chb.2013.01.005

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: understanding without words. *Journal of Experimental Psychology: General, 115,* 281–94. doi: 10.1037/0096-3445.115.3.281

Pouw, W. T., De Nooijer, J. A., Van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology, 5*: 359. doi: 10.3389/fpsyg.2014.00359

Pressley, G. M. (1976). Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology, 68,* 355. doi: 10.1037/0022-0663.68.3.355

Proverbio, A. M., & Riva, F. (2009). RP and N400 ERP components reflect semantic violations in visual processing of human actions. *Neuroscience Letters, 459,* 142-146. doi: 10.1016/j.neulet. 2009.05.012

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. doi: 10.1111/j.1460-9568.2005.03900.x

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7,* 226–231. doi: 10.1111/j.1467-9280.1996.tb00364.x

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422. doi: 10.1037/0033-2909.124.3.372

Renkl, A. (2011). Instruction based on examples. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 272-295). New York: Routledge.

Rimé, B. (1983). Nonverbal communication or nonverbal behavior? Towards a cognitive-motor theory of nonverbal behavior*. Current Issues in European Social Psychology, 1,* 85-141.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169-192. doi: 10.1146/annurev.neuro.27.070203.144230

Rocca, M. A., Falini, A., Comi, G., Scotti, G., & Filippi, M. (2008). The mirror-neuron system and handedness: A “right” world? *Human Brain Mapping, 29,* 1243–1254. doi: 10.1002/hbm0. 20462

Rohbanfard, H., & Proteau, L. (2011). Effects of the model’s handedness and observer’s viewpoint on observational learning. *Experimental Brain Research, 214,* 567-576. doi: 10.1007/s00221-011-2856-z.

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers’ word learning in a novel language. *Contemporary Educational Psychology, 38,* 109-117. doi:10.1016/j.cedpsych.2012.12.001.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nature Neuroscience, 4,* 546-550. doi: 10.1038/87510

Rueschemeyer, S. A., Van Rooij, D., Lindemann, O., Willems, R. M., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience, 22,* 1844-1851. doi: 10.1162/jocn.2009.21310

Rueschemeyer, S. A., Lindemann, O., Van Rooij, D., Van Dam, W. O., & Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology, 57,* 260-266. doi: 10.1027/1618-3169/a000031

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22,* 642-647. doi: 10.1111/j.1469-8986.1985.tb01661.x

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Memory and Cognition, 18,* 367-379. doi: 10.3758/BF03197126

Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly, 21,* 221-238. doi: 10.1080/10573560590949359

Salisbury, D. F. (2004). Semantic memory and verbal working memory correlates of N400 to subordinate homographs*. Brain and Cognition, 55,* 396-399. doi: 10.1016/j.bandc.2004.02.057

Salmon, N., & Pratt, H. (2002). A comparison of sentence-and discourse-level semantic processing: An ERP study. *Brain and Language, 83,* 367-383. doi: 10.1016/S0093-934X(02)00507-2

Saltz, E., & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning and Verbal Behavior, 20,* 322-332. doi: 10.1016/S0022-5371(81)90472-2

Sato, M., & Bergen, B. K. (2013). The case of the missing pronouns: Does mentally simulated perspective play a functional role in the comprehension of person? *Cognition, 127,* 361-374. doi: 10.1016/j.cognition.2013.02.004

Saxe, R., Jamal, N., & Powell, L. (2006). My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex, 16,* 178-182. doi: 10.1093/cercor/bhi095

Scharine, A. A., & McBeath, M. K. (2002). Right-handers and Americans favor turning to the right. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 44*, 248-256. doi: 10.1518/0018720024497916

Schmidt-Kassow, M., Kulka, A., Gunter, T. C., Rothermich, K., & Kotz, S. A. (2010). Exercising during learning improves vocabulary acquisition: Behavioral and ERP evidence. *Neuroscience Letters, 482,* 40-44. doi: 10.1016/j.neulet.2010.06.089

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide.* Psychology Software Incorporated.

Schrooten, W., & Vermeer, A. (1994). *Woorden in het basisonderwijs, 15.000 woorden aangeboden aan leerlingen* [Words in primary education, 15.000 words presented to pupils]. Tilburg: Tilburg University Press.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57,* 149–174. doi: 10.3102/00346543057002149

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*, 313–322. doi: 10.1037/0022-0663.77.3.313

Schwartz, R. M., & Raphael, T. E. (1985). Concept of definition: A key to improving students' vocabulary. *The Reading Teacher*, 198-205.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences, 12,* 411-417. doi: 10.1016/j.tics.2008.07.006

Shibata, H., Gyoba, J., & Suzuki, Y. (2009). Event-related potentials during the evaluation of the appropriateness of cooperative actions. *Neuroscience Letters, 452,* 189-193. doi: 10.1016/j.neulet.2009.01.042

Sirigu, A., & Duhamel, J. (2001). Motor and visual imagery as two complementary but neurally dissociable mental processes. *Journal of Cognitive Neuroscience, 13,* 910-919. doi: 10.1162/089892901753165827

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real–world events: An electrophysiological investigation. *Psychophysiology, 40,* 160-164. doi: 10.1111/1469-8986.00016

Smith, B. D., Stahl, N., & Neil, J. (1987). The effect of imagery instruction on vocabulary development. *Journal of College Reading and Learning, 20,* 131-137. doi: 10.1080/10790195.1987.10849909

Smyth, M. M., Pearson, N. A., & Pendleton, L. R. (1988). Movement and working memory: Patterns and positions in space. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 40A,* 497-514. doi: 10.1080/02724988843000041

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes, 27,* 665-681. doi: 10.1080/01690965.2011.573220

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227-267. doi: 10.1016/0010-0277(94)00647-4.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 121,* 153–156. doi: 10.1111/1467-9280.00326

Stevanoni, E., & Salmon, K. (2005). Giving memory a hand: Instructing children to gesture enhances their event recall. *Journal of Nonverbal Behavior, 29,* 217-233.doi: 10.1007/s10919-005-7721-y

Stins, J. F., Kadar, E. E., & Costall, A. (2001). A kinematic analysis of hand selection in a reaching task. *Laterality, 6,* 347-367. doi: 10.1080/713754421

Sumner, P., Nachev, P., Morris, P., Jackson, S. R., Kennard, C., & Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron 54,* 697–711. doi: 10.1016/j.neuron.2007.05.016

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634. doi: 10.1126/science.7777863

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press.

Taylor, L. J., & Zwaan, R. A. (2009). Action in cognition: The case of language. *Language and Cognition, 1, 45-58.* doi: 10.1515/LANGCOG.2009.003

Tellier, M. (2005). L'utilisation des gestes en classe de langue: comment évaluer leur effet sur la mémorisation du lexique? [The use of gestures in a language class: How to evaluate their effect on memorization of vocabulary?] In M. Billières, P. Gaillard, & N. Spanghero-Gaillard (Eds.), *Actes du Premier colloque international de Didactique Cognitive: DIDCOG2005Document of the first conference of cognitive didactics (proceedings on CD-Rom).* Toulouse, France: Université de Toulouse-Le Mirail.

Tellier, M. (2007). *How do teacher's gestures help young children in second language acquisition?* In Proceedings of the meeting of International Society of Gesture Studies, ISGS 2005: Interacting Bodies. France: Lyon.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture, 8,* 219–235. doi: HAL:http://hal.archives-ouvertes.fr/hal-00375251/en/

Tessari, A., Canessa, N., Ukmar, M., & Rumiati, R. I. (2007). Neuropsychological evidence for a strategic control of multiple routes in imitation. *Brain: A Journal of Neurology, 130,* 1111-1126. doi: 10.1093/brain/awm003

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17,* 273-281. doi: 10.1162/ 0898929053124965

Thomas, L. E., & Lleras, A. (2009). Swinging into thought: Directed movement guides insight in problem solving. *Psychonomic Bulletin & Review, 16,* 719-723. doi: 10.3758/PBR.16.4.719

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006) Rehearsal in serial memory for visual-spatial information: evidence from eye movements. *Psychonomic Bulletin & Review, 13,* 452–457. doi: 10.3758/BF03193869

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions*. Journal of Experimental Psychology: Human Perception and Performance, 24,* 830-846. doi: 10.1037/0096-1523.24.3.830.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology, 77,* 593–601. doi: 10.1037/h0026069

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 433-460.

Vacca, R. T., & Vacca, J. L. (2002). *Content area reading: Literacy and learning across the curriculum*. 6th ed. New York: Addison Wesley.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204. doi: 10.1016/S0361-476X(02)00007-3.

Valeri-Gold, M. (1994). *Making connections through reading and writing.* Connecticut: Wadsworth Publishing.

Van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657-671. doi: 10.1162/089892999563724

Van Dam, W. O., Rueschemeyer, S. A., & Bekkering, H. (2010). How specifically are action verbs represented in the neural motor system: An fMRI study. *NeuroImage, 53,* 1318-1325. doi: 10.1016/j.neuroimage.2010.06.071

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22,* 155-174. doi : 10.1007/s10648-010-9134-7

Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review, 21,* 21-30. doi: 10.1007/s10648-008-9094-3

Wassenaar, M., & Hagoort, P. (2007). Thematic role assignment in patients with Broca's aphasia: Sentence-picture matching electrified. *Neuropsychologia, 45,* 716-740. doi: 10.1016/ j.neuropsychologia.2006.08.016

Wellsby, M., & Pexman, P. M. (2014). Developing embodied cognition: insights from children's concepts and language processing. *Frontiers in Psychology, 5,* 1-10. doi: 10.3389/ fpsyg.2014.00506

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *American Journal of Psychology, 114,* 591–600. doi: 10.2307/1423612

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research, 13,* 363-375. doi: 10.1016/S0926-6410(01)00129-X

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in Psychology, 2*: 116. doi: 10.3389/fpsyg.2011.00116

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and language, 101,* 278-289. doi: 10.1016/j.bandl.2007.03.004
Willems, R. M., & Hagoort, P. (2009). Hand preference influences neural correlates of action observation. *Brain Research, 1269,* 90–104. doi: 10.1016/j.brainres.2009.02.057

Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Bodyspecific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science, 21*, 67–74. doi: 10.1177/0956797609354072

Willems, R. M., Labruna, L., D'Esposito, M., Ivry, R., & Casasanto, D. (2011). A functional role for the motor system in language understanding evidence from theta-burst transcranial magnetic stimulation. *Psychological Science, 22,* 849-854. doi: 10.1177/0956797611412387

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex, 17,* 2322-2333. doi: 10.1093/cercor/bhl141

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience, 20,* 1235-1249. doi: 10.1162/jocn.2008.20085

Willems, R. M., Toni, I., Hagoort, P. & Casasanto, D. (2009). Body-specific motor imagery of hand actions: Neural evidence from right- and left-handers. *Frontiers in Human Neuroscience, 3,* 39. doi: 10.3389/neuro.09.039.2009

Willems, R. M., Toni, I., Hagoort, P. & Casasanto, D. (2010). Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience, 22,* 10, 2387-2400. doi: 10.1162/jocn.2009.21386

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636. doi: 10.3758/BF03196322

Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: The SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *Journal of Cognition and Culture*, *5*, 165-190.

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

Zwaan, R. A. (in press). Situation models, mental simulations and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review.*

Zwaan, R.A. & Madden, C.J. (2005). Embodied Sentence Comprehension. In H.A.E.M. Pecher & R.A. Zwaan (Eds.), *The Grounding of Cognition: The rol of perception and action in memory, language and thinking* (pp. 224-245). Cambridge: Cambridge University Press.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE, 7,* e51382. Retrieved from http://www.plosone.org /article/info%3Adoi%2F10.1371%2Fjournal.pone.0051382.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135,* 1-11. doi: 10.1037/0096-3445.135.1.1

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

▪ Nederlandse samenvatting

Summary in Dutch

Achtergrond

In de studies beschreven in dit proefschrift werd onderzocht hoe motorische informatie het leren en onthouden van woorden kan beïnvloeden. Het leren van nieuwe woorden is een belangrijk onderdeel van het verwerven van een (nieuwe) taal (e.g., Folse, 2004), of dit nu de eerste, tweede of derde taal is. Het is dan ook van belang om effectieve instructietechnieken te vinden die het, soms moeizame, proces van woordleren kunnen verbeteren. In dit proefschrift werd onder andere onderzocht of en zo ja, hoe, activatie van het motorische systeem tijdens het leren van nieuwe woorden, leerprestaties kan verbeteren. Het idee dat motorische activiteit in de hersenen opgewekt door het observeren, imiteren of genereren van gebaren, invloed kan hebben op het leren van woorden komt onder andere voort uit de *embodied cognition* theorie.

Deze theorie stelt dat cognitieve processen, zoals taal en denken, niet los gezien kunnen worden van hoe ons lichaam functioneert (e.g., Barsalou 1999). Verder wordt aangenomen dat wanneer je een woord leest als 'schrijven', een mentale simulatie van deze handeling gecreëerd wordt. Hierbij wordt onder andere de eerdere motorische ervaring die je met die handeling hebt gehad, opnieuw geactiveerd. Bewijs voor dit idee komt onder andere vanuit onderzoek dat laat zien dat wanneer mensen woorden lezen of horen zoals 'schrijven', dit automatch leidt tot activatie in de motorische cortex van de hersenen. Dit betekent dat het taalsysteem en het motorsysteem van de hersenen elkaar kunnen beïnvloeden (e.g., Hauk, Johnsrude, & Pulvermüller, 2004). Vanuit dit *embodied cognition* perspectief kan verwacht worden dat door het activeren van het motor systeem (in de vorm van mentale simulaties), een rijkere representatie ontstaat van dat wat geleerd moet worden, omdat informatie in verschillende modaliteiten wordt aangeboden (zowel verbaal als motorisch). Op deze manier kan een dieper begrip ontstaan van het woord waarvan de betekenis geleerd moet worden. Vanuit dit perspectief wordt in dit proefschrift onderzocht, hoe het lichaam een rol kan spelen bij het beïnvloeden van leerprestaties.

De doelen van dit proefschrift waren (1) onderzoeken wat voor kinderen de beste manier is om gebaren te gebruiken tijdens het leren van nieuwe woorden in de eerste taal en nagaan of deze gebaren even effectief zijn voor kinderen met hogere en lagere taalvaardigheden en voor het leren van verschillende soorten werkwoorden; (2) onderzoeken of het gebruik van gebaren bij het leren van woorden in de eerste taal tot dezelfde resultaten leidt bij volwassenen en kinderen; (3) bestuderen of het leren en onthouden van object-manipulatie werkwoorden (bijvoorbeeld 'beitelen', 'boetseren') beïnvloed wordt door het zien van handelingen die niet overeenkomen met de mentale simulaties die automatisch gemaakt worden wanneer de definitie van een object-manipulatie werkwoord wordt gehoord; (4) verklaren waar de mogelijke effecten van mentale simulaties op taalverwerking vandaan komen. Hieronder wordt een samenvatting van de studies en de belangrijkste bevindingen in het licht van deze doelen gegeven.

In de studie die wordt gerapporteerd in **Hoofdstuk 2** werd onderzocht wat voor kinderen de beste manier is om gebaren te gebruiken tijdens het leren van nieuwe woorden in de eerste taal. Eerder onderzoek suggereert dat de leerprestatie kan verbeteren door tijdens het leren van woorden gebaren te observeren, te observeren en daarna te imiteren, of zelf een gebaar te verzinnen (e.g., Casale, 1985; Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011; Tellier, 2005). Een directe vergelijking tussen deze verschillende manieren waarop gebaren gebruikt kunnen worden tijdens het leren was echter nog nooit gemaakt. Er zijn wel studies die suggereren dat het actieve gebruik van het lichaam door gebaren te imiteren of gebaren te maken die de leerling zelf heeft verzonnen een sterker effect hebben op het leren van nieuwe woorden dan het alleen observeren van gebaren (e.g., Tellier, 2007). Basisschoolkinderen uit groep 5 leerden drie soorten nieuwe Nederlandse werkwoorden: abstracte (bijvoorbeeld 'afpoeieren', 'kniezen'), voortbeweging (bijvoorbeeld 'slenteren', 'paraderen') en object-manipulatie werkwoorden (bijvoorbeeld 'beitelen', 'boetseren'). De kinderen leerden deze werkwoorden door twee keer te luisteren naar een omschrijving van de betekenis van het woord. Tijdens de tweede keer dat ze de omschrijving van de betekenis hoorden, kregen ze ofwel (1) geen gebaren te zien, (2) observeerden ze een gebaar overeenkomstig met de betekenis van het woord (bijvoorbeeld voor het woord 'beitelen' zagen ze een slaande beweging in de richting van een gesloten hand), (3) observeerden en imiteerden ze het gebaar, (4) bedachten ze zelf een gebaar bij de omschrijving van de handeling en voerden dit uit.

De resultaten lieten zien dat voortbewegingswerkwoorden beter geleerd werden wanneer een gebaar geobserveerd werd dan wanneer alleen de definitie van het woord gegeven werd. In tegenstelling tot de verwachting werd de prestatie op het leren van voortbewegingswerkwoorden slechter wanneer kinderen een gebaar moesten imiteren of zelf een gebaar moesten verzinnen. Dit was niet het geval voor de object-manipulatie woorden en de abstracte woorden. Het leren van object-manipulatie woorden werd alleen beïnvloed door het gebruik van gebaren bij kinderen met hoge taalvaardigheden. Kinderen met lage taalvaardigheden leken eerder gehinderd te worden wanneer ze hun lichaam moesten gebruiken tijdens het leren (imitatie of zelf een gebaar verzinnen), vergeleken met wanneer ze geen gebaren zagen tijdens het leren. Een verklaring hiervoor kan zijn dat gebaren voor object-manipulatie werkwoorden cognitief complexer kunnen zijn (Cartmill, Beilock, & Goldin-Meadow, 2012), omdat het object (bijvoorbeeld de beitel) er zelf bij bedacht moet worden. Omdat voor kinderen met hoge taalvaardigheden, de belasting van het werkgeheugen tijdens het leren vermoedelijk lager is, hadden zij voldoende cognitieve capaciteit over om gebaren te imiteren en te maken en hadden zij daardoor wel profijt hiervan, in tegenstelling tot kinderen met lagere taalvaardigheden. Op het leren van abstracte woorden had het gebruik van gebaren tijdens het leren geen enkele invloed. Dit zou kunnen komen doordat abstracte woorden een minder sterke link met het motor systeem hebben dan werkwoorden die object-manipulatie of voortbeweging uitdrukken. De effecten die hierboven beschreven zijn, bestonden alleen voor scores op een test waarbij de kinderen werd gevraagd om de woordbetekenis die ze hadden geleerd zo letterlijk mogelijk op te schrijven. Wanneer de kinderen een taak moesten invullen waarin begrip van de nieuwe woorden gemeten werd, waren er geen effecten van het gebruik van gebaren tijdens het leren. Kortom, of het gebruik van gebaren tijdens het leren van werkwoorden effectief is voor basisschoolkinderen lijkt afhankelijk te zijn van het soort woord dat geleerd moet worden en van de taalvaardigheden van de kinderen.

In **Hoofdstuk 3** werd opnieuw de invloed van het gebruik van gebaren tijdens het leren van verschillende soorten nieuwe werkwoorden onderzocht. Deze keer lag de focus op het effect van het imiteren van gebaren op verschillende momenten. In Hoofdstuk 2 waren de positieve effecten van het imiteren van gebaren ten opzichte van het alleen observeren van gebaren niet overtuigend of zelfs afwezig, terwijl eerdere studies leken te suggereren dat het imiteren van gebaren leidt tot betere leerprestaties dan het enkel observeren van gebaren (Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011; Tellier, 2007; 2008). In deze studie is daarom een directe vergelijking gemaakt tussen het observeren van gebaren en het observeren en imiteren van gebaren tijdens het leren van werkwoorden, bij kinderen die een jaar ouder waren (groep 6) dan de kinderen uit Hoofdstuk 2. Daarnaast werd een factor onderzocht die wellicht een rol kan spelen in de effectiviteit van het imiteren versus het observeren van gebaren als gemeten middels toetsprestatie, namelijk het moment waarop een gebaar geïmiteerd wordt. De vraag die hier centraal stond was: Wat is het beste moment waarop een gebaar geïmiteerd kan worden? Is dit meteen nadat de omschrijving van het woord gehoord is en het bijbehorende gebaar geobserveerd is (dit zou kunnen helpen bij het *opslaan* in het geheugen), of is het effectiever om kinderen het eerder geobserveerde gebaar uit het geheugen *op te laten halen* en te laten uitvoeren net voordat de betekenis van het woord op een toets moet worden ingevuld? Of zou beide doen het meest effectief zijn? In het laatste geval is namelijk de context waarin het woord geleerd wordt en waarin de nieuwe kennis opgehaald wordt hetzelfde, wat een positief effect kan hebben op de leerprestaties (Tulving & Osler, 1968).

De resultaten lieten zien dat het imiteren van gebaren de toetsprestatie niet verbeterde ten opzichte van het enkel observeren van gebaren wanneer abstracte of voortbewegingswerkwoorden geleerd moesten worden. Echter, wanneer object-manipulatie woorden geleerd moesten worden, leidde het imiteren van gebaren tot betere toetsprestaties dan het observeren van een gebaar. Dit gold zowel voor het imiteren van een gebaar meteen na het horen van de omschrijving en het zien van het gebaar, als voor het imiteren van het gebaar net voor de toets. Werd het gebaar echter op beide momenten geïmiteerd, verbeterde de prestatie niet significant ten opzichte van het enkel observeren van het gebaar. Toen de kennis van de nieuw geleerde woorden na een week opnieuw werd getest (zonder dat er geïmiteerd werd), bleek dat de kinderen die een week eerder het gebaar net voor de toets imiteerden de meeste woorden hadden onthouden. Net als in de studie in Hoofdstuk 2, beperkte de verbeterde prestatie op zowel de eerste als tweede toets zich tot de toets waarin de letterlijke betekenis van het woord opgeschreven moest worden; op een toets van begrip van de woorden had het imiteren van gebaren geen effect. Kortom, het imiteren van een gebaar net voor het maken van een toets lijkt dus niet alleen het ophalen van de definitie die eerder is opgeslagen te vergemakkelijken, maar dit succesvol ophalen lijkt ook te leiden tot het beter onthouden van de betekenis van het nieuwe woord op de langere termijn.

De resultaten van de studies uit Hoofdstuk 2 en Hoofdstuk 3 zijn met elkaar in overeenstemming: Het positieve effect van het imiteren van gebaren op woordleren trad alleen op bij het leren van object-manipulatie werkwoorden (in Hoofdstuk 2 gold dit alleen voor kinderen met hoge taalvaardigheid, maar de kinderen in Hoofdstuk 3 waren een jaar ouder). Een mogelijke verklaring hiervoor zou kunnen liggen in de sterkere relatie tussen taal en handgebaren (Gentilucci & Corballis, 2006) dan tussen taal en andere gebaren. Een andere verklaring zou kunnen liggen in de doelgerichtheid van de handeling. Abstracte of voortbewegingswerkwoorden, bijvoorbeeld, zijn minder doelgericht dan object-manipulatie werkwoorden (de woorden 'vrezen' of 'slenteren', bijvoorbeeld, geven geen doel aan) en er is onderzoek dat laat zien dat handelingen moeilijker te imiteren zijn wanneer ze niet doelgericht zijn (Bekkering, Wohlschläger, & Gattis, 2000).

In de studie in **Hoofdstuk 4** werd het effect van het imiteren van gebaren op het leren van nieuwe object-manipulatie werkwoorden opnieuw onderzocht. In deze studie werd een vergelijking gemaakt tussen het effect op het leren bij basisschoolleerlingen (groep 6) en jongvolwassenen. Daarnaast werd in dit hoofdstuk ook onderzocht of de effecten van gebaren tijdens het leren beïnvloedt worden door de bekendheid van het object-manipulatie werkwoord waarvan de definitie geleerd moet worden. De verwachting was dat gebaren het leren meer zouden ondersteunen wanneer het woord dat geleerd moet worden compleet nieuw is, omdat gebaren hier extra (motorische) informatie bieden, wat kan helpen bij het opslaan van het nieuwe woord (bijvoorbeeld leren dat het woord 'hieuwen' betekent dat het anker van een schip binnengehaald wordt). Wanneer de betekenis van een object-manipulatie werkwoord geleerd moet worden waarvan zowel het woord als de handeling bekend zijn (bijvoorbeeld leren dat het woord 'typen' betekent dat woorden worden gevormd met behulp van een toetsenbord), en alleen de betekenis van het woord onthouden moet worden, kan dit betekenen dat de informatie uit de gebaren overbodig wordt. Uit eerder onderzoek blijkt dat overbodige informatie niet bijdraagt aan het leren, of het leren zelfs kan hinderen (Kalyuga & Sweller, 2014).

De resultaten lieten verrassend genoeg geen effect zien van imitatie tijdens het leren. Er was wel een interactie-effect tussen de leeftijdsgroep (kinderen – volwassenen) en taak (imitatie vs. alleen observatie), dat liet zien dat imitatie van gebaren het leren van kinderen niet verbeterde ten opzichte van observatie, terwijl imitatie bij volwassenen een negatief effect leek te hebben: zij leken beter te presteren wanneer de woorden geleerd werden door middel van het enkel observeren van gebaren. Dit gold zowel voor het leren van de betekenis van bekende als van nieuwe woorden; de effectiviteit van het imiteren van gebaren werd dus niet beïnvloed door de bekendheid van het woord waarvan de betekenis geleerd moest worden. Dat kinderen in deze studie niet profiteerden van het imiteren van gebaren in vergelijking met het observeren van gebaren, is verrassend in het licht van de resultaten uit Hoofdstuk 2 en 3. Een mogelijke verklaring is echter dat er onder de deelnemers aan de studie uit Hoofdstuk 4 een substantieel aantal kinderen lage taalvaardigheden had. Aangezien de resultaten in Hoofdstuk 2 al lieten zien dat vooral kinderen met hogere taalvaardigheden profijt lijken te hebben van het imiteren van gebaren, kan dit een mogelijke verklaring zijn voor waarom in deze studie geen effect van imitatie is gevonden.

In **Hoofdstuk 5** werd het effect van motoractivatie op woordleren op een heel andere manier bestudeerd. Wanneer mensen een omschrijving van een bekende handeling horen, maken ze automatisch een mentale simulatie van die handeling (alsof ze hem zelf uitvoeren; Casasanto, 2009). In de drie experimenten beschreven in dit hoofdstuk werd onderzocht of het zien van afbeeldingen die tegenstrijdig zijn met de mentale simulatie die de proefpersoon maakt op basis van de omschrijving van het woord, het woordleren zou beïnvloeden. In deze studie leerden volwassenen een aantal woorden uit een niet bestaande taal (bijvoorbeeld ‘otu’ betekent iets met twee handen kneden om het vorm te geven; ofwel boetseren, maar die ‘vertaling’ kregen ze niet te horen) in combinatie met plaatjes waarin een handeling werd afgebeeld vanuit een linkshandig of rechtshandig perspectief of met beide handen werd uitgevoerd. Alle plaatjes werden vanuit een eerste persoonsperspectief getoond. Omdat eerdere studies hebben laten zien dat het horen van de definitie van een object-manipulatie werkwoord of het zien van een plaatje van een handeling kan leiden tot lichaam-specifieke mentale simulaties, zal een rechtshandige zich automatisch een rechtshandige handeling voorstellen (Casasanto, 2009). Om deze reden kan voor een rechtshandige het zien van een plaatje waarin een linkshandige een handeling uitvoert een conflict creëren met de mentale simulatie die gecreëerd wordt na het horen van de definitie van het woord, waardoor het leren van het nieuwe woord gehinderd zou kunnen worden.

De resultaten toonden inderdaad aan, dat rechtshandigen meer definities leerden van woorden wanneer ze deze aangeboden kregen in combinatie met een plaatje waarin een handeling rechtshandig of met beide handen wordt uitgevoerd. Beide perspectieven komen overeen met de mentale simulatie, terwijl het linkshandige plaatje een conflict oplevert. Voor linkshandigen was er echter geen verschil in het aantal woordbetekenissen dat correct werd onthouden na het leren van woorden in combinatie met het zien van een afbeelding waarin een handeling met een linkerhand, rechterhand of beide handen werd uitgevoerd. Dit kan komen doordat linkshandigen in een rechtshandige wereld leven en daardoor vaker gedwongen worden hun rechterhand te gebruiken en daarom ook meer in aanraking komen met het rechtshandige perspectief. Zij zouden daarom vermoedelijk minder snel een conflict ervaren bij het zien van een rechtshandig plaatje.

In **Hoofdstuk 6** werden de grenzen van het effect uit hoofdstuk 5 opgezocht, door na te gaan of het effect afhankelijk is van observatieperspectief (eerste persoons- vs. derde persoonsperspectief). Afbeeldingen vanuit het derde persoonsperspectief lijken minder *embodied* te zijn dan afbeeldingen vanuit een eerste persoonsperspectief, dat wil zeggen: het eerste persoonsperspectief heeft een nauwere verbinding met het sensorisch motorische systeem dan het derde persoonsperspectief, dat nog verdere visuo-spatiele verwerking verlangt (Jackson, Meltzoff, & Decety, 2006). De effecten van derde persoonsperspectief afbeeldingen op het leren van woorden, kunnen daarom anders zijn. Naast het effect van derde persoons afbeeldingen is ook onderzocht of het zien van linkshandige plaatjes in eerste en derde persoonsperspectief het leren van rechtshandigen zou schaden in vergelijking met leren zonder afbeeldingen.

De resultaten van het eerste experiment lieten zien dat het leerresultaat van rechtshandigen niet negatief beïnvloed werd door het zien van een linkshandig derde persoonsperspectief in vergelijking met een rechtshandig derde persoonsperspectief. Het tweede experiment liet zien dat de leerprestatie niet verschilde wanneer woorden geleerd werden met en zonder afbeeldingen, ongeacht perspectief (d.w.z. het zien van een rechtshandig plaatje hielp niet en het zien van een linkshandig plaatje hinderde niet vergeleken met geen plaatje). Het effect uit Hoofdstuk 5 werd wel gerepliceerd voor het eerste persoonsperspectief: woorden met linkshandige eerste persoonsperspectief plaatjes werden minder goed geleerd dan woorden met rechtshandige plaatjes. Dat rechtshandigen wel beïnvloed worden door het leren van woorden in combinatie met een eerste persoons- maar niet met een derde persoonsperspectief, ligt vermoedelijk aan de eerder beschreven bevindingen dat het eerste persoonsperspectief een nauwere verbinding heeft met het sensorisch-motorische systeem. Het zien van een derde persoonsperspectief afbeelding tijdens het leren van nieuwe woorden lijkt daarom geen conflict te vormen met de mentale simulatie die gemaakt is op basis van het horen van de definitie van het woord.

In **Hoofdstuk 7** werd de aard van deze mentale simulaties verder onderzocht en werd de vraag gesteld of het negatieve effect van linkshandige plaatjes ook optreedt bij het simpelweg onthouden van bekende werkwoorden. In het eerste experiment moesten proefpersonen (jongvolwassenen) 20 woorden uit hun eerste taal onthouden. Ze kregen 5 seconden een object-manipulatie werkwoord op het scherm te zien en een afbeelding waarop deze handeling (vanuit een linkshandig of rechtshandig eerste persoonsperspectief) werd afgebeeld. Proefpersonen gingen ofwel meteen door naar het volgende item, of er was een pauze van twee seconde tussen de items, of proefpersonen moesten een beweging met hun rechterhand maken om verder te kunnen naar het volgende woord. De resultaten lieten zien dat wanneer er een pauze van twee seconden was tussen de items, rechtshandigen minder woorden onthielden die gepaard gingen met een linkshandige afbeelding dan woorden die gepaard gingen met een rechtshandige afbeelding. Als er geen pauze tussen de items zat of wanneer de rechtshandigen een armbeweging met de rechterhand moesten maken was er geen verschil in hoeveel woorden de proefpersonen hadden onthouden.

In het tweede experiment werd de vraag onderzocht of het mogelijke effect van het combineren van een woord met een linkshandig of rechtshandig eerste persoonsperspectief plaatje op het geheugen kan liggen aan de manier waarop volwassenen naar deze plaatjes kijken. Met behulp van eye-tracking apparatuur werden de oogbewegingen van proefpersonen geregistreerd. De oogbewegingsdata toonden aan dat tijdens de 5 seconden dat het plaatje en woord in beeld waren er geen verschillen waren in het kijkgedrag naar de links- en rechtshandige perspectief plaatjes. Wanneer alleen naar de eerste fixatie (een pauze van de oogbeweging op een specifiek gebied in het visuele veld) werd gekeken leken deze iets langer te duren bij de linkshandige plaatjes, maar dit verschil was niet significant. Interessanter waren de oogbewegingen tijdens het twee seconden durende interval tussen de woorden, waarin het scherm leeg was. Ondanks het lege scherm keken proefpersonen langer naar de plek waar de items eerder gepresenteerd waren dan naar andere gebieden, wat vooral gebeurde wanneer ze net een plaatje met een rechtshandig perspectief hadden gezien. Dit kan erop duiden dat de proefpersonen op dat moment het woord dat ze moesten onthouden, actief herhaalden of ophaalden uit het geheugen (e.g., Martarelli & Mast, 2013). Omdat de proefpersonen langer naar de lege plek keken waar de afbeelding eerder was gepresenteerd wanneer het een rechtshandig-perspectief plaatje betrof en de proefpersonen meer woorden onthielden die gekoppeld waren aan een rechtshandig-perspectief plaatje, kunnen de langere blikken naar de lege ruimte bij de rechtshandig-perspectief plaatjes erop duiden dat het ophalen van het woord dat onthouden moest worden succesvoller was (ook in ander onderzoek is namelijk de link tussen geheugen voor stimuli en kijken naar de locatie waar die stimulus voorheen zichtbaar was, gelegd: e.g., Johansson & Johansson, 2014; Kent & Lamberts, 2008. Kortom, ook al was er slechts een klein effect van mentale simulaties op het leren (Hoofdstuk 5 en 6) en onthouden (Hoofdstuk 7) van woorden, dit effect lijkt wel robuust te zijn bij in rechtshandigen.

In **Hoofstuk 8** werden mentale simulaties nogmaals onderzocht om meer te weten te komen over de aard van mentale simulaties en de invloed op taalverwerking. Dit werd gedaan door proefpersonen een zin te laten horen waarin een handeling werd beschreven (bijvoorbeeld 'jij snijdt de tomaat') en hen daarbij afbeeldingen te laten zien van een congruente handeling (tomaat snijden) of incongruente handeling (t-shirt strijken) vanuit een rechtshandig of linkshandig eerste persoonsperspectief. Aan de proefpersonen werd gevraagd om zo snel mogelijk met hun dominante hand aan te geven of het plaatje overeenkomt met de inhoud van de zin, waarbij reactietijden gemeten werden. Daarnaast werd tijdens het uitvoeren van deze taak een EEG (elektro-encefalogram) gemaakt. Onderzocht werd of het zien van incongruente plaatjes die niet overeenkomen met de inhoud van de zin, tot een groter N400-effect leiden. De N400 is een component die 400 milliseconden na het presenteren van de stimulus (het plaatje) in het EEG-signaal gevonden kan worden in de vorm van een negatiefgaande piek. Deze component reflecteert het neurale mechanisme van semantische integratie in een (zins)context. Wanneer verbale en visuele informatie succesvol wordt geïntegreerd, is slechts een relatief kleine piek te zien. Als de twee bronnen van informatie niet goed geïntegreerd kunnen worden, is er een grotere negatieve piek te zien, 400 ms nadat beide bronnen van informatie beschikbaar werden. De centrale vraag was of het zien van een afbeelding die niet overeenkomt met de handvoorkeur ook in de N400 teruggevonden kan worden. Deze studie zou daarom inzicht kunnen bieden in of mentale simulaties een rol spelen in het vroege proces van het integreren van informatie uit de verbale en visuele modaliteit.

De resultaten lieten zien dat hoewel er een N400 effect was voor inhoud, wat betekent dat het moeilijker is om een plaatje te integreren in de context van de zin wanneer het plaatje niet overeenkomt met de inhoud van de zin, dit effect niet gemoduleerd werd door het perspectief (links of rechts) van het plaatje. Het is daarom ook onwaarschijnlijk dat het integreren van een plaatje in de context van de zin beïnvloed wordt door lichaam-specifieke mentale simulaties. De resultaten van de reactietijdentaak, die op een later moment ook met een groep linkshandigen is uitgevoerd zonder EEG meting, lieten zien dat rechtshandigen langzamer reageerden op linkshandige plaatjes, terwijl linkshandigen even snel op links- als op rechtshandige plaatjes reageerden. Het zien van een afbeelding die niet overeenkomt met de mentale simulatie die op basis van handvoorkeur gemaakt wordt, lijkt daarom wel een effect te hebben op gedragstaken, zoals reactietijden, geheugen en leren (Hoofdstuk, 5, 6, 7, en 8), maar op basis van deze studie lijkt er geen indicatie te zijn dat mentale simulaties een rol spelen in een vroeg integratie proces zoals gemeten met het N400 effect.

In **Hoofdstuk 9** werden de praktische en theoretische implicaties van de studies besproken. Om met het laatste te beginnen: De resultaten van het onderzoek beschreven in dit proefschrift komen niet overeen met de traditionele kijk op cognitie (e.g., Fodor, 1983; Turing, 1950), waarin motorische informatie geen rol zou moeten spelen in een cognitief proces, zoals leren. De bevindingen dat het observeren of imiteren van gebaren het leren van nieuwe woorden kan verbeteren (Hoofdstuk 2 en 3) zou binnen deze theorie daarom niet verwacht worden. De resultaten beschreven in dit proefschrift zijn meer in lijn met een *embodied cognition* perspectief, dat stelt dat cognitieve processen zoals taal en denken niet los gezien kunnen worden van hoe ons lichaam functioneert (e.g. Barsalou, 1999; Glenberg & Gallese, 2011). Echter, de bevindingen komen zeker niet overeen met een *puur* embodied cognition standpunt waarin wordt geclaimd dat cognitie geen amodale representaties nodig heeft en volledig gegrond is in het sensorisch-motorisch systeem (e.g., Wilson, 2002). Als dit het geval zou zijn, dan zou het gebruik van gebaren ook het leren van abstracte woorden moeten verbeteren, maar dit was niet het geval (Hoofdstuk 2 en 3). Ook zouden we dan verwachten dat mentale simulaties altijd taalverwerking zou beïnvloedden, wat ook niet het geval was (Hoofdstuk 6, 7, 8). De resultaten zijn meer in lijn met een minder strikte variant van *embodied cognition,* waarin gesteld wordt dat cognitie gebaseerd is op zowel sensorisch-motorische als amodale informatie, (e.g., Taylor & Zwaan, 2009; Zwaan, 2014; Zwaan, in press; Zwaan & Madden, 2005), en sensorisch-motorische informatie, amodale informatie kan verrijken en kan zorgen voor een dieper niveau van verwerking.

Een nieuwere theorie stelt echter dat er niet meer gekeken moet worden naar of we abstracte symbolen of *grounded* symbolen (symbolen die gegrond zijn in de sensorimotorische ervaring) gebruiken voor cognitieve processen, maar naar wanneer we welke symbolen gebruiken. Deze *pluralist view*, stelt dat cognitieve processen gebruik maken van zowel *grounded* als abstracte symbolen (e.g., Dale, Dietrich, & Chemero, 2009; Dove, 2009; Zwaan, 2014). Het gaat er dan niet zozeer meer om óf we mentale simulaties maken, maar wanneer en onder welke omstandigheden. Zo werd in Hoofdstuk 6 beschreven dat het zien van afbeeldingen met een linkshandig perspectief, het leren van rechtshandigen negatief kan beïnvloeden als deze afbeeldingen vanuit een eerste persoonsperspectief getoond worden, waarschijnlijk omdat dit conflicteert met hun rechtshandige eerste persoonsperspectief mentale simulatie die ze maakten op basis van het horen van de betekenis van het woord, terwijl er geen conflict optreedt wanneer handelingen vanuit een derde persoonsperspectief getoond worden.

Tot slot, zijn de resultaten van deze studies ook interessant in het licht van theorieën over multimedia leren (e.g., Mayer, 2014) en multi-sensorische input (e.g., Shams & Seitz, 2008). Binnen beide theorieën wordt gesteld het gebruik van congruente multimodale informatie (i.e., auditief/verbaal, visueel/pictorieel, motorisch) tot beter leren leidt dan alleen informatie uit één modaliteit. Als we kijken naar de bevindingen uit Hoofdstuk 2 en 3, zijn deze bevindingen moeilijk te verklaren vanuit een multisensorische theorie. Want waarom zou het effect van gebaren op leren dan verschillen tussen verschillende woordsoorten? Het niet vinden van een effect voor abstracte werkwoorden kan vanuit een meer pluralistisch embodied cognition account verklaard worden, door de geringere koppeling met het motorische systeem van deze woorden. De effecten van gebaren komen daarom waarschijnlijk niet alleen door het toevoegen van een extra (i.e. motorische) modaliteit, maar heeft vermoedelijk meer te maken met de sensorisch-motorische informatie die de gebaren bieden. Kortom, de studies beschreven in dit proefschrift leveren bewijs voor een visie op cognitie waarin mentale simulaties gebruikt worden, maar niet altijd nodig zijn voor taalbegrip, waarbij het van de context afhangt of mentale simulaties cognitieve processen, zoals leren, kunnen beïnvloeden.

De resultaten die in dit proefschrift gerapporteerd worden, hebben ook relevantie voor de onderwijspraktijk, bijvoorbeeld voor het gebruik van gebaren in de les. Het leren van nieuwe woorden is een belangrijk deel van het leren van een taal, of dat nu een eerste of een tweede taal is. Op basis van eerder onderzoek leeft het idee dat het gebruiken van gebaren tijdens het leren van nieuwe woorden, dit proces zou kunnen verbeteren. De resultaten van dit proefschrift suggereren echter dat het gebruik van gebaren tijdens woord leren lang niet altijd nuttig is, maar afhangt van het type woorden en leerlingkenmerken. Het leren van abstracte werkwoorden verbeterde in mijn studies bijvoorbeeld niet, wanneer gebaren gebruikt werden. Het leren van concrete werkwoorden (object-manipulatie en voortbewegingswerkwoorden) leek wel te verbeteren wanneer gebaren werden gebruikt, waarbij voortbewegingswerkwoorden beter geleerd werden wanneer gebaren alleen geobserveerd werden, terwijl object-manipulatiewerkwoorden beter geleerd leken te worden wanneer gebaren ook geïmiteerd werden. ‘Leken’, want het effect van het imiteren van gebaren tijdens het leren van objectmanipulatie werkwoorden hing echter wel af van de taalvaardigheden van de leerling (die samenhangen met de werkgeheugencapaciteit). Tot slot lijkt het gebruik van gebaren het leren van nieuwe woorden alleen te stimuleren op een taak waarbij de betekenis van het nieuwe woord letterlijk opgeschreven moet worden. Gebaren lijken daarom op basis van de resultaten van dit proefschrift de prestatie op een begripstest niet te verbeteren. Er moet dan ook nog verder onderzoek gedaan worden naar de omstandigheden waaronder observeren, observeren en imiteren of zelf genereren van gebaren wel of niet effectief is voor het leren van nieuwe woorden.

De studies beschreven in dit proefschrift laten ook zien dat het gebruiken van plaatjes die de woordbetekenis weergeven het woordleren niet significant verbeteren (maar ook niet verslechteren) in vergelijking met het woordleren zonder plaatjes, in elk geval bij volwassenen. Wanneer er echter wel plaatjes worden getoond tijdens het leren, lijkt het raadzaam om hierbij wel rekening te houden met de handvoorkeur van de leerling / student. Het gros van de leerlingen is rechtshandig en voor hen hebben plaatjes waarop een linkshandige de handeling uitvoert, een negatief effect op hoeveel nieuwe woorden worden onthouden, of hoeveel definities geleerd worden door rechtshandigen, vergeleken met plaatjes vanuit een rechtshandig perspectief.

▪ Dankwoord

Acknowledgements in Dutch

Een proefschrift schrijven doe je niet alleen en ik ben daarom ook dank verschuldigd aan verschillende mensen die dit proefschrift mogelijk hebben gemaakt.

Ten eerste mijn begeleiders: Tamara van Gog, Fred Paas en Rolf Zwaan, om onze auteursvolgorde maar aan te houden.

Tamara, dankjewel voor je fijne begeleiding, steun, adviezen, hulp en onvermoeibare luisterende oor. Ze hebben mijn promotietraject heel aangenaam gemaakt. Ik had me geen betere dagelijks begeleider kunnen wensen. Dankjewel ook voor alle gezelligheid en leuke gesprekken tijdens congressen, feestjes en borrels. Ik hoop dat dit niet het einde daarvan is.

Fred, bedankt voor de altijd snelle feedback op mijn stukken. Waar je ook was ter wereld, ik kon altijd rekenen op een reactie in een fractie van seconden. Dat heeft me zeer geholpen. Dankjewel voor je bereidheid om altijd te helpen en bij te dragen aan onze studies.

Rolf, bedankt voor de goede discussies tijdens labmeetings en de duidelijke feedback op mijn stukken, die mijn werk een stuk makkelijker maakten. Bedankt ook voor het af en toe binnen lopen om gewoon even te vragen hoe het was en voor je waardevolle adviezen die me altijd weer vooruit hielpen.

Zonder collega's zou mijn promotietraject een stuk minder aangenaam geweest zijn en zou dit proefschrift van mindere kwaliteit zijn. Veel dank gaat daarom uit naar hen. In het bijzonder wil ik iedereen bedanken binnen de Brain & Cognition groep en alle aanwezigen bij onze wekelijkse labmeetings (in Belmondo, Student hotel, en op een blauwe maandag in een zeer dubieus senioren centrum): Lisa, Karen, Nicole, Jan, Lysanne, Wim, Anna en Anita, bedankt voor de feedback op mijn studies, voor alle momenten dat ik even binnen kon lopen met een vraag en voor alle leuke gesprekken, pubquizavondjes en andere gezelligheid. Liselotte, dankjewel voor de fijne samenwerking bij ons paper en de EEG practica. Verder wil ik ook alle pubgroupleden bedanken: Margot, Marit, Kim, Martine, Vincent, Mario, Noortje, Tim, Gerdien, Margina, Gertjan, Steven, Marloes, Daniel, Lydia, Lisette, Sofie, Peter, bedankt voor alle discussies en het kritisch lezen en reviewen van een aantal van mijn stukken en natuurlijk voor de bezoekjes aan de *pub* daarna. Tot slot mijn kamergenoten, Sabine, met jou heb ik het langst een kamer gedeeld. Ik vond het fijn om een kamer te delen met iemand die er dezelfde ideeën over promoveren op nahield als ik. Stijn, Charly (heel eventjes, maar toch), Andrea en Danyang, thanks for all the nice talks we have had over the years, and for the opportunity to whine to each other when things weren't going our way. Tot slot, Margot en Marit, bedankt dat jullie mijn paranimfen willen zijn, en achter mij willen staan tijdens mijn verdediging. Ik vind het heel leuk dat we onze etentjes ook buiten de EUR gewoon doorzetten.

Een onderzoek kan niet uitgevoerd worden zonder proefpersonen en daarom wil ik hierbij alle proefpersonen die aan mijn onderzoeken hebben meegewerkt bedanken. Ook de kinderen die over de jaren heen altijd vol enthousiasme mijn taakjes hebben ingevuld wil ik bij deze bedanken, zonder jullie was dit proefschrift er niet gekomen. Om het onderzoek bij al die proefpersonen af te kunnen nemen, was ik vaak te vinden in het Erasmus Behavioral Lab en

daarom wil ik ook Christiaan, Marcel en Gerrit Jan bedanken voor hun hulp en steun. Door jullie liepen mijn onderzoeken altijd gesmeerd.
Graag wil ik ook de leden van de promotiecommissie bedanken. De leden van de kleine commissie: dr. Katinka Dijkstra, Prof. dr. Jan van Strien en Prof. dr. Harold Bekkering en de leden van de grote commissie. Bedankt dat jullie met mij van gedachten willen wisselen over mijn proefschrift.

Tot slot wil ik ook hen bedanken, die niet direct iets met dit proefschrift te maken hebben gehad maar mijn traject wel een stuk aangenamer hebben gemaakt. Mijn vrienden, bedankt voor alle gezellige borrels en etentjes na het werk, die even alle onderzoeksmechanismen in mijn hoofd stil legden. Mijn familie, klein maar fijn zullen we maar zeggen. Papa, zonder jou was ik niet zo ver gekomen. Dankjewel dat je het zo lang met me vol hebt gehouden en me de vrijheid hebt gegeven om de wijde wereld in te trekken.

Lieve Michael, registered partner in crime, dankjewel dat je mijn leven zoveel leuker maakt, me uitdaagt, er bent als ik dat nodig heb, altijd gek met me wilt doen en heel de wereld met me over trekt. Ik hoop dat we nog lang van ons mooie leventje mogen genieten.

▪ Publications

In press

De Nooijer, J. A., Gootjes, L., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). Picturing meaning: An ERP study on the integration of left- or right-handed first-person perspective pictures into a sentence context. *Language, Cognition and Neuroscience.*

De Nooijer, J. A., & Willems, R. M. (2015). What can we learn about cognition from studying handedness? Insights from cognitive neuroscience. In N. Hagemann, B. Strauss, C. MacMahon & F. Loffing (Eds.). *Laterality in Sports: Theories and Applications.* San Diego, California: Academic Press, Elsevier.

Hald, L. A., De Nooijer, J. A., Van Gog, T., & Bekkering, H. (2015). Optimizing word learning via links to perceptual and motoric experience. *Educational Psychology Review.*

2014

Pouw, W. T. J. L., De Nooijer, J. A., Van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology, 5*, 539. doi: 10.3389/fpsyg.2014.00359.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2014). Words in action: Using gestures to improve verb learning in primary school children. *Gesture, 14,* 47–70. doi 10.1075/gest.14.1.03noo.

2013

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica*, *144*, 173-179. doi: 10.1016/j.actpsy.2013.05.013.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives. *Psychological Science, 24,* 2515-2521. doi: 0.1016/j.actpsy. 2013.05.013.

Submitted

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *Effects of pictures with left- or right-handed first-person and third-person perspectives on learning of object-manipulation words.* Manuscript submitted for publication.

De Nooijer, J. A., Van Wermeskerken, M., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *Recalling the right words: Handedness effects on recalling verbs using left or right-handed first-person perspective pictures.* Manuscript submitted for publication.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2015). *The effects of gesture observation versus imitation on children's and adult's learning of novel and known action verbs.* Manuscript submitted for publication.

Kock, R., Van Gog, T., & De Nooijer, J. A. (2015). *Learning to solve the tower of hanoi problem by self-modeling vs. modeling examples.* Manuscript submitted for publication.

Van Schalkwijk, F. J., Benjamins, J. S., Migliorati, F., Van Someren, E. J. W., De Nooijer, J. A., Van Gog, T., & Van der Werf, Y. D. (2015). *The role of sleep timing in children's observational learning.* Manuscript submitted for publication.

Presentations

Paper

De Nooijer, J. A., Gootjes, L., Van Gog, T., Paas, F., & Zwaan, R. A. (August, 2014). *Picturing meaning: An ERP study on the integration of left or right-handed first-person perspective pictures into a sentence context.* Paper presented at ESLP 2014, the 7th Annual Conference of Embodied and Situated Language Processing, Rotterdam, the Netherlands.

De Nooijer, J. A., Van Gog, T., Van Wermeskerken, M., Paas, F., & Zwaan, R. A. (July, 2014). *Recalling the right words: Handedness effects on recalling action verbs learned with left or right-handed action pictures.* Paper presented at ISGS 6, the 6th Conference of the International Society for Gesture Studies, San Diego, United States of America.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (August, 2013). *Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance.* Paper presented at EARLI 2013, the 15th Biennial Conference of the European Association for Research on Learning and Instruction, Munich, Germany.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (June, 2013). *When left Is not right: Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives.* Paper presented at SARMAC 2013, the 10th Biennial Conference of the Society for Applied Research in Memory and Cognition, Rotterdam, The Netherlands.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (July, 2012). *Words in action: Using gestures to improve verb learning in primary school children.* Paper presented at ISGS 5, the 5th Conference of the International Society for Gesture Studies, Lund, Sweden.

Poster

De Nooijer, J. A, Van Gog, T., Paas, F., & Zwaan, R. A. (December, 2014). Leren een handje helpen. Hoe ons lichaam onze leerprestaties beïnvloedt. Poster presented at the 'Breinproductendag' in Utrecht, The Netherlands.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (June, 2013). *When left is not right: Handedness effects on learning object-manipulation words using pictures with left or right-handed first-person perspectives.* Poster presented at ESLP 2013, the 6th Annual Conference of Embodied and Situated Language Processing, Potsdam, Germany.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (May, 2013). *Words in action: Using gestures to improve verb learning in primary school children.* Poster presented at AERA 2013, the Annual Conference of the American Educational Research Association, San Francisco, United States of America.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (August, 2012). *Words in action: Using gestures to improve verb learning in primary school children.* Poster presented at ESLP 2012, the 5th Annual Conference of Embodied and Situated Language Processing, Newcastle upon Tyne, United Kingdom.

Invited Talks

De Nooijer, J. A. (March, 2015). Leren een handje helpen: Hoe ons lichaam onze leerprestaties beïnvloedt. Masterclass presented at TinQwise, Hilversum.

De Nooijer, J. A. (January, 2015). *Wetenschapsles.* Primary school de Egantelier, Soest.

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (December, 2014). *Leren een handje helpen. Hoe ons lichaam onze leerprestaties beïnvloedt.* Breinproductendag, National Initiative Brain and Cognition, Utrecht.

De Nooijer, J. A. (November, 2014). *Leren doe je met je handen! Hoe gebaren helpen bij het leren van nieuwe woorden.* Kennisdag Cubiss, Tilburg.

De Nooijer, J. A., Post, L. S., & Van Huik, B. (April, 2012). *Leren met gebaren op de basisschool; Taal en natuurkundige concepten.* Symposium Universiteit Leiden.

De Nooijer, J. A. (April, 2012). *Nieuwe woorden leer je met je handen: Een onderzoek naar de rol van embodied cognition in woordleren.* CED-groep, Rotterdam.

De Nooijer, J. A. (November, 2011). *Words in action.* Brein in Beeld Symposium, NEMO Amsterdam.

▪ Curriculum Vitae

Jacqueline Angelique de Nooijer was born in Vlissingen on June 10th, 1986. She completed secondary education (Gymnasium) in 2004 at the Christelijke Scholengemeenschap Walcheren in Middelburg. She obtained her Bachelor's degree in Language and Culture Studies with a major in language and cognition and a minor in English language and culture in 2007 from Utrecht University (Cum Laude). She then obtained a Master's degree in Language and Speech Pathology at the Radboud University Nijmegen in 2009 (Bene Meritum), during which time she also worked as a research assistant at the Max Planck Institute for Psycholinguistics in Nijmegen. In 2009 she entered an Erasmus Mundus European Master in Clinical Linguistics for which she studied at the University of Milano-Bicocca (Italy), the Rijksuniversiteit Groningen (the Netherlands) and the University of Potsdam (Germany), for which she received a Master's degree in 2011 (Classification A, Excellent). In 2011 she started working as a Ph.D. candidate at the Institute for Psychology at the Erasmus University Rotterdam that resulted in this dissertation. The studies in this PhD project focused on the role of motor activation on language processing and language learning and were supervised by Prof. dr. T. A. J. M van Gog, Prof. dr. F. Paas and Prof. dr. R. A. Zwaan. In addition to conducting research, Jacqueline supervised several Bachelor and Master theses and was involved in teaching a number of undergraduate and graduate courses. In addition she reviewed articles for international journals, was a member of several organizing committees and was selected as a participant for KNAW's Faces of Science.