CrossMark

REVIEW ARTICLE

Optimizing Word Learning via Links to Perceptual and Motoric Experience

Lea A. Hald[1] · **Jacqueline de Nooijer**[2] ·
Tamara van Gog[2] · **Harold Bekkering**[1]

Abstract The aim of this review is to consider how current vocabulary training methods could be optimized by considering recent scientific insights in how the brain represents conceptual knowledge. We outline the findings from several methods of vocabulary training. In each case, we consider how taking an embodied cognition perspective could impact word learning. The evidence we review suggests that vocabulary training methods can be optimized by making a congruent link between the words to learn and one's own perceptual and motoric experiences. In particular, we suggest that motoric information about the meaning of a word could be incorporated into more standard vocabulary training methods. Finally, we consider the impact an embodied cognitive perspective may have on other characteristics of word learning, such as individual differences in learning and variations in learning different types of words, for example words from different word classes and words in different contexts.

Keywords Word learning · Multimedia learning · Action observation · Motor simulation · Embodied cognition · Gesture

✉ Lea A. Hald
l.hald@donders.ru.nl; lea.hald@gmail.com

Jacqueline de Nooijer
j.a.denooijer@fsw.eur.nl

Tamara van Gog
vangog@fsw.eur.nl

Harold Bekkering
h.bekkering@donders.ru.nl

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

[2] Institute of Psychology, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Springer

Interest in vocabulary development in school-aged children has regained importance in reading instruction and reading research over the last decade (see Pearson et al. 2007). Pearson et al. (2007) suggest that this trend is due, in part, to reports that vocabulary has not received enough focus lately (e.g., National Institute of Child Health and Human Development 2000). The aim of this review is to show how vocabulary teaching methods could be further improved by taking into account findings from recent neurocognitive research that shows that word representation and word learning are grounded in perception and action. It is often said that word learning works best when a link can be established between what the learner already knows and what they are trying to learn (e.g., Beck et al. 2002; Mezynski 1983; Nagy and Herman 1987; Stahl and Nagy 2006). What the learner already knows is usually defined in terms of other words that the learner knows or situations that may be familiar to the learner (e.g., Beck et al. 2002). We suggest that another, potentially more powerful, approach would be to define "what the learner already knows" in terms of their previous motoric and perceptual experiences. We outline different routes for word learning in the brain and in particular focus on an embodied cognition perspective (e.g., Glenberg and Gallese 2012) that emphasizes the grounding of word meaning in motoric and perceptual experiences via the sensorimotor system in the brain.

Why Motoric and Perceptual Experience?

Evidence suggests that word meaning is represented in the brain in a manner that reflects real-world experience with words' referents, specifically with the motoric and perceptual experience that occurred in combination with those words (Barsalou 2008; Fischer and Zwaan 2008; Glenberg and Kaschak 2002; Glenberg and Gallese 2012; Martin 2007; Willems and Hagoort 2007; Zwaan and Taylor 2006). Theories of cognition proposing an important role of perceptual and motor experience for grounding conceptual meaning are typically captured by the term embodied cognition. Although there is not a unified embodied theory of cognition (Glenberg 2010; Pezzulo et al. 2012; Wilson 2002), relying on perceptual and motor experience to ground meaning is central. For example, our concept of *birds* is based on the collection of bodily states (perceptual and motoric) we have experienced with birds. Broadly speaking, our understanding of the word *bird* arises from partially re-activated (simulated) previous bodily states. This reactivation or simulation is not considered to be a conscious mental image; instead, simulation is more like a record of previous neural states (see Barsalou 1999 for a general theory; see Glenberg and Kaschak 2002; Glenberg and Gallese 2012; Taylor and Zwaan 2009, for theories related more to language and embodied cognition). The idea that motor and perceptual experiences are the basis for word meaning and conceptual representation has important consequences for how we understand and learn language.

When it comes to concepts, there are many action concepts that have a strong motor component. For example, we use our motor experience to distinguish the difference between words like *saunter* and *walk*. Similarly, some concepts become almost impossible to describe without relying on motor information, such as *backhand tennis swing*. These examples point to the importance of motoric experience for understanding words. However, action does not work alone; evidence suggests a tight link between action and perception (e.g., Friston and Kiebel 2009).

Action is the means by which we induce most perceptual changes. When we perceive the world, we move our eyes, rotate our head, and move our body to induce new scenes.

Moreover, research suggests that movement is necessary to interpret the sensations that are available to us (see Clark 2013; Friston and Kiebel 2009 for a review). Essentially, the idea is that we continuously predict or model the sensory inputs from our body and the environment around us. The discrepancy or mismatch between the predicted sensory input and currently experienced information (i.e., prediction error) results in changing either the model (perception) or changing the environment (action; Friston and Kiebel 2009, see also Glenberg and Gallese 2012 for related claims about action systems being used to derive predictions about the incoming environment). Therefore, the role of action is crucial for understanding whether changes in sensation are produced by us (we are moving and the environment is stable) or whether the changes in the sensation are produced by a changing environment (while we are stable). In other words, our bodily actions make things happen, which in turn creates meaning.

Research suggests that understanding concepts relies directly on bodily action. For example, experiments where young children interact with objects and people in naturalistic settings indicate that early category learning is constrained by action and by hand actions on objects (Yoshida and Smith 2008). However, the process of learning new concepts appears to be specifically responsive to the link between perception and action (see Yu and Smith 2012). For example, when infants start to independently execute manual exploration, they begin to show more advanced object segregation abilities (Needham 2000) and begin to attend more to different sensory properties of objects (Eppler 1995). The importance of action for multimodal object representations remains important for adults (see Van Elk et al. 2014 for an extensive review of supporting evidence). The importance of action for understanding is also evidenced in research using action to aid reading comprehension in school-aged children. For example, a technique called *moved by reading* involves children simulating text they read by moving toys or images on a computer screen (see Glenberg 2011 for a summary). The action-based simulation leads to improved comprehension (Glenberg et al. 2011).

The idea that action improves comprehension is not a new one. The entire Montessori movement is based on the general idea that moving and thinking are highly related (e.g., Montessori 1966), and research in cognitive neuroscience is increasingly uncovering evidence to support this relationship. For instance, the link between action and language has been demonstrated by studies suggesting that hearing action words, such as *kick*, activate an area of the brain that overlaps with the area activated when actually executing the denoted action (the brain's motor and premotor cortex; e.g., Hauk et al. 2004; Pulvermüller et al. 2005; Tettamanti et al. 2005; for extensive reviews of the link between language and action, see Gallese 2007; Glenberg and Gallese 2012; Pulvermüller 2005; Rizzolatti and Craighero 2004).

The link between sensory perception and language is similarly robust. There are many concepts that have a strong perceptual component to them. For instance, we use our gustatory experience to distinguish concepts like *sour*, *tart*, and *piquant* but rely more on our tactile and visual experience to distinguish concepts like *cotton*, *feathery*, and *silken*. Brain-imaging research supports this idea. Increased activation in visual areas of the brain has been found when participants made a decision about visual properties of words (e.g., *which things are green*) and an increase in the somatosensory cortex when people had to make decisions about tactile properties of words (e.g., *which things are soft*; Goldberg et al. 2006).

In sum, there is considerable evidence that motoric and perceptual experiences are the basis for learning and representing concepts. The significance of this for understanding word meaning is captured by the embodied cognition framework. The evidence we review suggests that both L1 and L2 vocabulary training methods can be optimized by explicitly linking the

words to learn to one's own congruent motoric and perceptual experiences. Enhancing this link may lead to deeper understanding of word meaning and better transfer of meaning across texts.

Standard Vocabulary Training Methods

Probably, the most widely used way to teach new vocabulary is to simply provide a definition along with the target vocabulary. While this seems straightforward enough, an early meta-analysis of word learning studies suggests that definitions alone do not lead to the best learning results (Stahl and Fairbanks 1986). Instead, the meta-analysis found that providing a sentence(s) using the word in context, along with a definition, led to the best learning. Interestingly, these early findings are easily explained by the embodied cognition framework. According to the framework, skilled reading involves creating an embodied simulation from the context (Glenberg 2011). Therefore, providing a sentence using the target word in context should induce embodied encoding more than a definition would, which in turn should enhance vocabulary acquisition.

Another straightforward way to train vocabulary is with drill-and-practice methods. Drill-and-practice techniques often involve many repetitions of the same information (typically definitions) in order to build associations. However, extra exposures to definitions via drill-and-practice methods do not lead to better learning (as shown by a meta-analysis; Stahl and Fairbanks 1986). What appears to affect learning is the type of information students are exposed to rather than the number of exposures. In studies that control for the amount of time spent on training (i.e., using drill-and-practice methods), it has been shown that mixing a definition with the use of the word in context leads to better learning than just providing definitions (i.e., Jenkins and Dixon 1983; Stahl 1983).

There are many ways to provide context in order to improve the link between what the learner already knows and the to-be-learned word. For example, there are many verbal association methods, such as teaching words in semantically related groups. Using semantically related groups to teach new words has been shown to be effective when the semantically related words are graphically displayed with related information (e.g., Carr and Mazur-Stewart 1988). For example, concept wheels (Vacca and Vacca 2002), semantic word maps (Schwartz and Raphael 1985), and semantic feature analysis (Johnson and Pearson 1978) all use graphical displays to aid students in making a link between the word to learn and words they already know. Without discussing each of these techniques in extensive detail, an example of one will be provided to elucidate the discussion.

Semantic word mapping (Schwartz and Raphael 1985) relies on background knowledge of the student, but rather than just using words, it also creates a visual display that links related concepts with students' background knowledge (see Fig. 1).

In general, all of the above semantic relationship methods, as well as other methods that are aimed at enriching the type of information students are exposed to, lead to better learning outcomes than simple definitions alone (see Beck et al. 2002). The extra information helps improve the connection between what the learner already knows and the to-be-learned word (Beck et al. 2002). Although not typically framed in terms of motoric and perceptual experience, the example of semantic word mapping can easily be explained in terms of improving word learning via links with previous perceptual experience. Learners can link their previous visual experience of seeing shells and snails to abalone. Essentially, these methods provide a link between what the learner already knows via perceptual experience
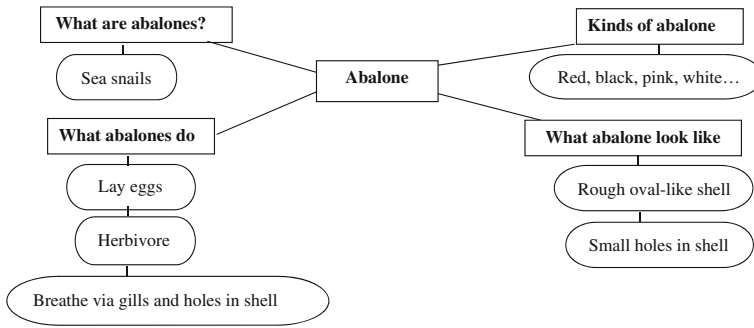
Fig. 1 An example of a semantic word map where concepts about the meaning are visually linked together (Note: This is an original example figure)

and word meaning. Although the particular example illustrated here does not obviously link to previous motoric experience, upon further examination, there is a motoric component to the semantic information. Abalone is described above as "looking" rough and oval with small holes. However, the visual experience of roughness, "ovalness," and having small holes is something that can easily be linked to actually picking up and feeling a shell on a beach or in an aquarium. Arguably previous experience of picking up and feeling a rough, oval shell will lead to a much richer understanding than simply viewing a rough shell. And according to the embodied cognition framework, describing the motoric and perceptual information in this way enriches the representation of the word meaning, which explains why semantic relationship methods improve word learning. Recent evidence supports this claim. Across four experiments, students had physical experience with science concepts (e.g., angular momentum; Kontra et al. in press). They found improved performance on quiz scores when students had the physical experience related to the science concept. Furthermore, brain imaging indicated that improved performance was due to sensorimotor brain regions being active when students later reasoned about angular momentum (Kontra et al. in press).

Imagery Vocabulary Training Methods

In a different tradition of vocabulary training methods, providing a link between what the learner already knows and the to-be-learned word is provided through the use of imagery. The use of imagery in vocabulary training is not typically described in terms of motoric and sensory experiences. However, as is argued for the traditional vocabulary training methods above, the effectiveness of imagery training can also benefit from considering previous motoric and sensory experience.

Imagery training can be accomplished through a variety of methods. One such method is the keyword method (Atkinson 1975). The aim of the keyword method is to improve the associative link between the word-to-be-learned and the definition. Instead of focusing on rote memorizing of the definition of a word, a keyword is used to create an interactive image that can connect the word to the definition. Keyword methods are often used for learning a second language. For example, learning that the French word "pain" means bread could be done by focusing on the fact that "pain" sounds like "pan" in English. The learner could then generate an image of a loaf of bread on a pan (Pressley et al. 1987). Overall, the keyword method leads

to high learning outcomes (e.g., Pressley, et al. 1987; Stahl and Nagy 2006). The keyword method encourages connections between conscious images of previous experiences of the learner and the meaning of the word.

This type of mental imagery has been shown to lead to beneficial effects on a variety of reading and vocabulary tasks (e.g., Anderson and Kulhavy 1972; Bull and Wittrock 1973; Pressley 1976; Sadoski and Paivio 1994). For example, Bull and Wittrock (1973) used pictures and mental imagery to teach definitions of 16 nouns that were drawn from a seventh-grade spelling list to 87 fifth graders. Specifically, students received the words with a verbal definition only, a verbal definition plus illustration, or an imagery condition in which students were given a verbal definition and asked to create their own accompanying image. In the imagery condition, the students were asked to draw their own illustration of the definitions. At the test, 1 week later, the students in the imagery group outperformed the definition only group. Although a trend was seen for the definition plus illustration group to outperform the definition only group, this effect did not reach statistical significance. More recent imagery training research has also indicated a positive impact on vocabulary acquisition in school-aged children (e.g., Cohen and Johnson 2011). Specifically Cohen and Johnson (2011) found that children who received imagery training for words (again using mental imagery and drawing) outperformed children who only received words paired with definitions. However, no difference was seen between the imagery training and the words paired with picture conditions. Overall results suggest that mentally generating an image followed by drawing the image leads to better learning than pairing words with definitions and may also lead to better learning than being provided with an image. The difference in learning could be due to the extra engagement with the materials that children have when self-generating an image (e.g., dual coding, Cohen and Johnson 2011). Embodied cognition theories make similar predictions in terms of effectiveness of imagery but for a somewhat different reason. According to embodied cognition theories, generating an image should be more effective than a definition due to the additional sensory and especially motor experience that would co-occur while generating the image (mental or drawn), and this process would later lead to a richer mental simulation of the meaning of the word.

Related to using imagery, the cognitive linguistics field has developed imagery methods to help teach the meaning of words that do not have one clear literal meaning. For instance, cognitive linguistics has used imagery methods for teaching metaphorical meanings, idioms, as well as other more difficult words to learn, such as prepositions or phrasal verbs (e.g., *follow through with*). Most of this research focuses on teaching these meanings in L2. Although the specifics of the method may vary, the common technique is to begin with having students think about the literal or basic meaning of a word. For example, to think about what *mushroom* means related to mushrooms on a pizza. Following this, the students are asked to make an educated guess at the meaning of the metaphorical use of the word. For example, what *mushroom* means in *The number of grocery stores in the area has mushroomed*. Although the literal meanings can be defined verbally, students are asked to represent the metaphorical meanings with either the use of imagery (Szczepaniak and Lew 2011), enactment (Lindstromberg and Boers 2005) or pictures and drawings (Boers et al. 2009), which have been shown to lead to better learning (e.g., Lindstromberg and Boers 2005; see Boers 2013 for a recent review). In sum, this group of training methods combines verbal information for a literal meaning with some type of image or enactment to help students learn a non-literal meaning.

The use of imagery for word learning has often been explained in a manner comparable to embodied cognition. The Dual Coding Theory (DCT; Paivio 1971; 1986; 2007; Sadoski 2005,

2009; Sadoski and Paivio 1994, 2001, 2004; Sadoski et al. 1991) of vocabulary development explains cognition through the interaction of two mental codes: the verbal code and the nonverbal code. The verbal code deals with language in several sensory modalities, including speech and writing. The nonverbal code, primarily via mental imagery, deals with the representation and processing of objects, events, and situations in all sensory modalities (see Sadoski 2009). Essentially all memory, meaning, and knowledge are a result of the representation, processing, and interaction between these two codes. Although the DCT is not intended to explicitly specify the brain processes involved in constructing meaning, the DCT can account for the majority of the findings on imagery and word learning outlined here. Furthermore, recent versions of DCT have incorporated aspects of embodied cognition theories as well (see Sadoski 2009).

In terms of the embodied cognition perspective, these imagery techniques can be likened to simulating or re-activating prior perceptions and linking them to the word meaning. Interestingly, imagery techniques appear to promote the creation of a conscious link to prior perceptual or motoric experience, whereas simulation is not usually conscious. However, the underlying ideas of imagery and simulation are highly related. Furthermore, many theories of embodied cognition have proposed something like mental imagery to re-activate previous motoric and sensory experiences (e.g., Gallese and Lakoff 2005). Likewise, research with adults on action words supports this proposal. Actively imagining action words activates the primary and premotor cortex in the brain (e.g., Willems et al. 2010). The beneficial effect of providing imagery training is likely due to the link that is made to previous perceptual experiences.

If one accepts the idea that imagery training methods improve learning at least in part due to the link made between word meaning and previous perceptual experiences (see Miyashita 1995 for a short review; for recent evidence see Koziol et al. 2011; Schendan and Ganis 2012), one might still question how much imagery training has to do with motoric experience. Although in practice many examples using imagery tend to promote connections between prior *perceptual* experiences and word meaning, there is no principled reason for this. These connections could just as easily be based on previous action experiences. Related to this point, an early study comparing the use of motor enactment, visual imagery, and verbal statements for remembering sentences showed that young children and adults can remember both nouns and verbs in the sentences better after motoric enactment than verbal rehearsal or visual imagery (focused on visual information about the words; Saltz and Donnenwerth-Nolan 1981). Although this study is not about learning word meaning per se, it does show how powerful motoric information can be for memory processes. As is argued above for semantic relationship methods, examples of imagery methods that feel very perceptual in nature also tend to have a motoric component to them. Take for example learning the two words discussed previously, "mushrooming" and "pain." In the *mushrooming* example, it is quite likely that people imagine a neighborhood with many grocery stores and that this image involves them looking around or walking around the neighborhood and imaging new grocery stores appearing (i.e., mushrooming) at least to some degree. If it is the case that linking meaning to previous perceptual and motoric experience leads to greater learning, then it should follow that asking learners to explicitly incorporate motoric and rich perceptual information into their imagery would lead to improved learning. With learning the meaning of the French word "pain" by generating an image of a loaf of bread on a pan could involve students imagining themselves holding the pan with a loaf of bread or removing the loaf of bread from the pan. The embodied cognition account would predict that adding this motoric element to the

previously mainly perceptual image would enhance learning. To date, there are no experiments that were designed to address this possibility with imagery. Clearly, this remains an empirical question that needs to be answered. However, as will be discussed shortly, pieces of evidence from research where learners are actually asked to perform meaning-related movements, such as enactment, imitation, or gestures, suggest that the link to motoric experience is a powerful learning tool. Before turning to research on the link between actually moving and word learning, there is an additional way that the link between previous experience and word meanings have been made, namely using multimedia tools.

Multimedia Vocabulary Training Methods

Multimedia tools are yet another option for enhancing the link between previous motoric and perceptual experience and word meaning. In multimedia learning materials, the presentation of written or spoken verbal information is combined with static (e.g., photos, diagrams) or dynamic (video, animation) pictorial information (Mayer 2014). While multimedia tools often include the use of tablets or computers to present the information, books that include illustrative pictures are also multimedia. Multimedia therefore can involve more than one modality, but it need not do so. One can liken the use of visual aids, typically illustrations, to imagery, although as noted above, imagery sometimes produces better learning results than illustrations. Nonetheless, research exploring the use of verbal plus visual information, such as illustrations, suggests that the additional visual information can also promote better word learning (e.g., Smith et al. 1987).

Because the majority of multimedia techniques for word learning involve using static illustrations along with text (either presented on paper or by a computer), this review will begin with what is known about the use of illustrations before discussing other types of multimedia materials, such as auditory materials and animations. Several studies have found evidence that pairing an illustration with text improves word learning more than text alone. For example, Smith et al. (1987) had undergraduate students learn words that were unknown to them. The words were presented to them in one of three between subject conditions: a definition, a definition plus the word used in a sentence context, or a definition, a context sentence and a picture illustrating the meaning of the word. After a 2-week delay, students' performance indicated significant improvement when they received the definition, context sentence, and picture compared to the definition alone. There was no significant difference between the definition plus context sentence and the other two conditions.

Similarly, improved word learning has been found for immediate and delayed tests when students received illustrations along with a definition and context sentence compared to definition plus context sentence alone (see Valeri-Gold 1994). Additional support for the use of pictures in word learning comes from McGregor et al. (2007). Using a between subjects design, they gave vocabulary lessons that included pictures and names of unfamiliar object referents. In particular, they varied the number of exposures as well as the richness of the semantic information given (pictures plus definitions). Learning was measured via a definition task and a picture-naming task. Results revealed an interaction between the number of exposures to the word and the semantic information given, namely, the higher the frequency of exposure, the greater immediate benefit for the definition task, with a delayed benefit for picture naming. Furthermore, higher frequency of exposures helped learning most (as measured by the definition task) when they were presented with definitions paired with pictures.

However, using multimedia does not necessarily lead to better word learning (e.g., Bull and Wittrock 1973). Some evidence suggests that the spatial layout of written text and illustrations can lead to processing difficulties and high cognitive load, which may hamper learning (see Sweller et al. 2011). That is, when written text (e.g., the word plus verbal definition or a text in which the new words are used) and illustration are presented physically separate (i.e., the picture above, beneath or next to the text), this results in *split-attention* (Ayres and Sweller 2014): students have to divide their attention over the two (or more) information sources and need to mentally integrate the information from these sources. Split-attention problems are dependent on the complexity of the materials (which in turn depends partially on prior knowledge of the learner; Sweller et al. 2011) and the amount of visual search required to locate corresponding information across different information sources. When extensive visual search is required, learners tend to have problems integrating the information, which can lead to excessive working memory load and in turn decreased learning. Ways to resolve the detrimental effects of split-attention are cueing corresponding elements in text and picture (De Koning et al. 2009; Van Gog 2014), physically integrating text and picture (Ayres and Sweller 2014), or using multiple modalities, that is, presenting the text as spoken instead of written text (Low and Sweller 2014). It should be noted that in general, split-attention problems are primarily found when the multimedia included two sources of visual information (picture and text), rather than a mixed modality presentation (spoken text and pictures). It is not just dual representations that lead to split-attention, but it appears to be related to multiple sources of visual information.

An example to illustrate how split-attention across multiple information sources hampers word learning comes from Acha (2009) who presented third and fourth grade Spanish speakers learning English as a second language with an English story that contained words to be learned. Each word was paired with several types of information: a visual text translation of the word, a picture illustrating the word, or the simultaneous presentation of text and picture. Recall of word translations was best when children had only received text translations compared to pictures or pictures and text. The authors suggest that this result is due to limited working memory capacity of the children. Indeed, people with lower working memory capacity, lower cognitive ability, or lower prior knowledge can be expected to be more sensitive to the detrimental effects of the high cognitive load caused by split-attention. Other studies have also shown that the effectiveness of simultaneous presentation of pictures with written text appears to be very sensitive to individual differences. For instance, in multimedia learning in general, it was found that in comparison with learners with low ability, learners with high verbal and visual ability appear to benefit from simultaneous presentation (Mayer and Sims 1994). With regard to word learning with multimedia in particular, English speakers learning German as a second language who had low cognitive ability were shown to learn new words best when either verbal or picture information was presented separately rather than simultaneously (Plass et al. 2003). Related to this, it has been demonstrated that when adults are given a choice as to how they would like the material to be presented, the presentation of words and pictures simultaneously led to greater learning than presenting only words or only pictures. Considering the research on split-attention and working memory load, we argue that multimedia can be beneficial for all learners when you present the two pieces of information in the right manner for that learner. For instance, high verbal and visual learners may benefit more from simultaneous cross-modal information when the information is either easy to integrate or cued, while low verbal learners may benefit more from sequential presentation of information, since this can lead to less cognitive load or split-attention problems (e.g., Low and Sweller 2014; Plass et al. 2003).

Like the use of imagery techniques, the benefit of multimedia methods has been described in terms of dual coding (Sadoski 2005). The crucial claim of DCT in relationship to multimedia training is that the more one can activate both the verbal and non-verbal code while learning a new word, the better the word will be learned. Illustrations can activate the non-verbal code and the text the verbal code, hence greater word learning. Although the DCT incorporates aspects of embodied cognition when explaining word learning, more explicit theories of embodied cognition go beyond DCT in outlining the brain processes involved (e.g., Glenberg and Gallese 2012; Willems, and Hagoort 2007) and in outlining how meaning is constructed in the brain (e.g., Barsalou 1999; Zwaan, and Taylor 2006). In describing how meaning is constructed in the brain, these embodied theories go beyond DCT in adding the importance of motoric information as well as sensory information beyond visual information (e.g., Glenberg and Kaschak 2002; Glenberg and Gallese 2012; Pulvermüller and Fadiga 2010). According to the embodied cognition perspective, making a congruent link between the to-be-learned words and the perceptions and actions that the learner has previous experience with will improve learning. Most vocabulary studies have investigated effects of providing verbal information along with illustrations, primarily via the visual domain. However, depending upon the type of multimedia employed, other sensory and even motoric information could be conveyed with illustrations. For example, animations can be used to illustrate motor information, auditory information, such as related sounds could be played on a computer, and even tactile or gustatory information could be used to illustrate meaning via objects related to the word to be learned.

While both still pictures and dynamic visualizations such as videos or animations may promote connections to previous perceptual experiences, dynamic visualizations may provide a richer link to both perceptual and motor experience (see Gazzola and Keysers 2009; Proverbio et al. 2009 for examples of motor activation during observation of static versus dynamic pictures). How an object looks, and what one can do with it, can become much clearer when you see it being used dynamically. For example, one's understanding of what a screwdriver is would be much richer after watching a how a screwdriver is used. Information about how heavy something is, what substance it might be made of, how you should hold it, and what action you perform with it can be demonstrated in a dynamic animation in a way that could promote connections to a range of previous perceptual experiences. One prediction that falls out of considering other sensory as well as motoric experience is that use of meaningful dynamic multimedia video may lead to greater word learning.

Dynamic Multimedia Vocabulary Training Methods

Using illustrations to aid word learning is not new, but the use of dynamic visualizations such as animations or videos for word learning is a relatively recent development. Some of the early research exploring the effectiveness of watching dynamic visualizations for learning in general was fairly disappointing.

In a review of the results of many studies comparing learning from animations compared to still pictures, Tversky et al. (2002) concluded that there was no difference between learning with animations compared to still illustrations (see also Hegarty and Kriz 2007). The lack of a beneficial effect of dynamic visualizations has been mainly attributed to cognitive load, caused by the transience of information (Sweller et al. 2011): if students do not attend to the right information at the right time, it may have disappeared and is no longer available for processing. Measures to counteract

negative effects of transience, such as cueing (De Koning et al. 2009; Van Gog 2014) or segmentation (Spanjers et al. 2010), can improve the effectiveness of dynamic visualizations. Furthermore, in their meta-analysis, Höffler and Leutner (2007) found that overall dynamic visualizations had a small positive effect compared to static visualizations.

With regard to word learning, early research with adults learning German as a second language that used both still pictures as well as short videos suggests that word learning is best when information about the word meaning is presented visually (via pictures or video) and verbally, rather than just verbally (Plass et al. 1998, 2003). For example, for the German word *Hubschrauber* (helicopter), a 7-s video clip of a helicopter was displayed; however, they did not actually design their study to compare still pictures versus animations. Instead, the comparison of interest was nonverbal versus verbal information, especially in relationship to the individual ability in verbal and visual information as tested separately. The benefit of the visual information (still picture or animation) depended highly on individual differences of the learner (e.g., Plass et al. 2003).

A recent study explored the effect of watching video versus listening to books on vocabulary learning (Silverman 2013). They found no difference between using books or videos on children's learning. The verbal ability of the children did not play a role in their results. Similarly, the effect was the same regardless of whether learning was assessed with receptive or expressive tests. Though when videos were shown repeatedly, vocabulary learning improved. However, this effect was only seen when the vocabulary assessment was expressive. In contrast, Smeets and Bus (2014) explored vocabulary improvement by assigning children to one of four reading conditions: static e-books, animated e-books, interactive animated e-books, and a control condition. Children in the control condition played nonliterary related computer games. The results indicated that gains in vocabulary were greatest when children read interactive animated e-books followed by animated e-books and then static e-books as compared to the control condition.

One area where dynamic visualizations have been shown to consistently be more effective for learning is visualizations of human movement tasks (Höffler and Leutner 2007; Van Gog et al. 2009). It has been proposed that this is due to the activation of the mirror neuron system when processing animations that include motor actions: when observing an action, the same neural circuits are activated that would be active when one performs an action oneself (Rizzolatti and Craighero 2004). In other words, viewing actions activates an automatic process of embodied simulation, which might reduce working memory load and lead to the beneficial effects on learning (Van Gog et al. 2009).

Theories of embodied cognition would suggest that the involvement of sensorimotor experience could also be marshalled to enhance learning of non-motor tasks, such as language, from dynamic visualizations. To apply the idea that observing movement by others will aid word learning, Hald et al. (2015) explored the effectiveness of dynamic visualizations on verb learning in children. In this study, Dutch primary school children (aged 7–8 years old) received one of three different animation training conditions in combination with the verbs they were asked to learn (e.g., *chisel*). In one condition, they watched a congruent animation whereby an avatar illustrates the specific movement(s) needed to perform that action (verb) on a particular object correctly. For example, the avatar held a chisel and performed a chiseling action on a large stone. A second condition entailed a congruent still picture whereby a screenshot of the correct animation was displayed. The third condition showed an incongruent animation whereby the same avatar made an incorrect movement in relationship to the meaning of the

verb. For example, holding a chisel, the avatar moved the chisel around in circles above the stone. In all three conditions, the children heard a context sentence using the verb correctly that was presented simultaneously with the animation/still picture (e.g., *Anna is chiseling a beautiful statue.*).

Results suggested that the congruent animation, by nature of depicting an action dynamically, improved learning compared to the still picture. In particular, compared to children who received still pictures, the children that saw congruent animations performed better on an expressive test of understanding (Hald et al. 2015).

A follow-up study (Hald et al. 2015) varied the information in the context sentence that was presented and found that children learned the greatest number of words when they saw a congruent animation paired with a context sentence that focused on the goal of the action (e.g., chiseling a statue) rather than the actual movement the avatar made (e.g., moving the hands/chisel up and down). The importance of focusing on the goal of an action compared to the actual kinematic movement in order to understand an action is a significant area of research in itself, beyond the scope of this review (see Ondobaka et al. 2012). However, the results suggest that although videos can aid word learning, it is important that the goal of the action depicted in the animation is clear.

Gesture Observation

Another area in which dynamic videos have been shown to be beneficial in word learning is when they included iconic gestures. Iconic gestures are those that have a perceptual relation to concrete entities or actions. For example, for the word *bridge*, an iconic gesture would be to move one's hands from left to right in an arc, like a bridge. Evidence that observing these gestures aids word learning comes from Tellier (2005). In this study where 5-year-olds were presented with a video of a person who only pronounced the words or additionally presented them with iconic gestures congruent with the meaning of the words they had to learn, children performed better on a free recall task when they had been presented with gesture and verbal information, compared to when they learned words without gestures (Tellier 2005). Word learning with gesture observation also seems to be effective in situations outside a lab setting. When parents of infants are trained to increase their verbal plus iconic gesture input to their infants, those infants have stronger receptive language and expressive language at 24 months than children that did not receive this input from their parents (Goodwyn et al. 2000). However, the effect of gesture observation seems to be dependent on the type of word that has to be learned. In a recent study, in which 8- to 9-year-olds learned novel abstract (e.g., to dismiss), locomotion (e.g., to stride), and object manipulation verbs (e.g., to chisel), gesture observation did not improve learning for abstract verbs, but it did for locomotion verbs. Learning of object manipulation verbs only seemed to improve under gesture observation for children with high verbal skills (De Nooijer et al. in press).

In sum, there seem to be two (not mutually exclusive) possible explanations for the beneficial effects of observing dynamic actions or gestures on learning: (1) the observation of actions or gestures leads to a richer mental representation of the learning material, and (2) because motor information is automatically processed, working memory load is reduced and learning is enhanced. Again, since the cognitive load imposed by a task is partially determined by the amount of prior knowledge and individual characteristics (e.g., working memory capacity, ability) of the learner, research on the effects of observing actions or gestures in

dynamic visualizations of words should consider effects of or interactions with individual differences.

The research on dynamic animations suggests that your motor system appears to help you understand what you see. Because action observation activates the motor system to a lesser extent than action execution (Gazzola and Keysers 2009), an interesting question is whether *performing* movements related to the meaning of the word itself while learning word meanings would also lead to greater word learning.

Use of Gesture Imitation in Vocabulary Training Methods

Along with observing action or gesture in videos or animations, much of the research on the use of gestures in videos for word learning has involved the imitation of gestures. Although there are indications that mere action observation facilitates learning, as described above, action execution has also been found to enhance memory traces as compared to receiving only verbal information (e.g., Engelkamp and Cohen 1991; Dijkstra and Kaschak 2006). Learning might therefore improve even more when the learner's own motor system is engaged, via action imitation. This benefit of gesture imitation over observation has also been found within the word-learning domain. For example, children can remember more items from a list with concrete nouns and descriptive adjectives from a second language when imitating gestures that accompany the words than when only observing these gestures (Tellier 2008). Furthermore, when adults are instructed to learn expressions from a second language with emblematic gestures (i.e., nonverbal acts which have a direct verbal translation and are well known within groups or cultures, Ekman and Friesen 1969), the expressions are learned better with emblematic gestures that were imitated by the participants, than when words were learned by only repeating the sentence without seeing the gesture (Allen 1995).

The effectiveness of gestures in word learning appears to depend on the characteristics of the learner (Rowe et al. 2013). Preschoolers learned artificial words for familiar objects via only the English translation of the word, or in combination with a picture congruent with the meaning of the word, or in combination with a congruent iconic gesture. Children were tested on immediate comprehension and comprehension after 1 week. Characteristics such as language ability, language background, and gender played a role in the effectiveness of these nonverbal aids. For example, girls performed significantly better than boys when learning words paired with pictures. In addition, the researchers found an interaction between the learning method, language background, and language ability, showing that these factors have an effect on whether or not nonverbal aids are useful in word learning. Using nonverbal aids in a group with high language abilities does not improve performance, while children with low verbal abilities benefit most from the use of gestures during word learning (Rowe et al. 2013).

However, these studies applied gesture imitation during encoding, immediately after observation of the gesture. The effectiveness of gesture imitation might be dependent on the moment at which the gesture is imitated. In a recent study, 9–11-year-olds learned new verbs (i.e., abstract, locomotion, or object manipulation verbs) in their first language with either observation of a congruent gesture, observation plus imitation during encoding (i.e., directly after hearing the definition of the word that had to be learned), during retrieval (i.e., during test-taking), or at both of these times. The gestures could be considered pantomimes as all actions were depicted as realistically as possible. On both an immediate and a delayed posttest after 1 week, children's word learning benefitted from gesture imitation during retrieval, while

imitation during encoding was only more beneficial than merely observing the gesture at an immediate posttest. This was, however, only true for object manipulation verbs. For locomotion and abstract verbs, gesture imitation did not improve performance more than gesture observation (De Nooijer et al. 2013). This study, therefore, suggests that gesture imitation in word learning might be applied during retrieval instead of during encoding the new word. It also suggests that the effectiveness of gesture imitation is dependent on the type of verb that has to be learned. These results are in line with a similar study on learning different types of verbs in combination with gesture observation and imitation. In this study, only performance on object manipulation verbs seemed to benefit from gesture imitation in 8- to 9-year-olds with high verbal skills (De Nooijer et al. in press).

A question that arises while describing the effects of gestures during word learning is whether the gestures used during learning have to be meaningful or whether any kind of action will do? The studies discussed above have mainly investigated word learning with the use of meaningful gestures. It might, however, be the case that mere motor activation will improve learning, because of the close link between action and language that has been proposed by theories of embodied cognition (e.g., Fischer and Zwaan 2008). According to this view, motor activation (or motor resonance) grounds the meaning of words and therefore enhances language comprehension, which could in turn improve learning. One study in which native speakers of German learned French nouns and verbs while being physically active (i.e., cycling) or passive is in line with this suggestion. Participants in the physically active group scored higher on vocabulary tests than participants in the physically passive condition, suggesting that mere motor activation can enhance word learning performance (Schmidt-Kassow et al. 2010). This suggests that any larger movements will help with learning, presumably due to motor activation. Nevertheless, when it comes to gesture, which typically involves less physical movements than cycling, the gesture movements appear to only be effective for learning if they are meaningful. For example, in a study where German native speakers learned words from an artificial language in combination with observing and imitating meaningful iconic gestures or non-meaningful gestures, only the use of meaningful gestures led to better memory performance (Macedonia, Müller, & Macedonia et al. 2011; see also Kelly et al. 2009).

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, before discussing the implications of this in more detail, we consider a related finding on action performance during learning, known as the enactment effect (see review by Engelkamp 1998; Nilsson 2000).

Use of Enactment in Vocabulary Training Methods

We define enactment here as self-generated actions, without the presence of an example (e.g., observation of a gesture by someone else). In the 1960s, the use of enactment was already incorporated in a teaching method whereby students had to perform actions to commands that were given by the teacher. The enactment was thought to enhance understanding and memory for vocabulary items in a foreign language (Asher and Price 1967). However, this technique never became a success because there was no empirical evidence to support its efficacy (Macedonia and von Kriegstein 2012).

Since then, many studies have investigated the role of self-generated action on memory. These studies 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 and Kaschak 2006; Engelkamp et al. 1994). For example, children were taught the meaning of new word using a motor imaging procedure. In this procedure, the children were given the verbal definition of the word and were asked to pantomime the meaning of a novel abstract verb, adjective, or noun. The pantomimes were first imagined by each child independently. Following that the most common pantomime for each word meaning was used for additional practice with all children. This pantomime method was compared to a standard verbal definition approach and the Manzo's subjective approach to vocabulary (Manzo 1982). The subjective approach involves a verbal definition with verbal examples followed by asking students in a classroom to "think of some personal images or experiences which you can associate or picture with it." If the students in the class do not suggest anything, then the teacher might suggest an association. The students learned the most words when they used pantomime (Casale 1985).

Similarly, in a longitudinal study, students learned words from an artificial language corpus 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 is improved when adults are cued during recollection with gestures that they made for the to be remembered words earlier compared to when they were not cued with gestures or with someone else's gesture (Frick-Horbury 2002). Nonetheless, the effect of self-generated action might not always be larger than the effect of action observation. When participants were asked to read phrases containing action verbs (e.g., to peel a potato) or were additionally required to watch a gesture (e.g., observing someone peeling a potato) or produce a mime (e.g., pretend to peel a potato themselves), a facilitative effect on memory was found for both action observation and enactment. However, scores in the condition where a gesture had to be observed were not different from scores in which a mime was produced, suggesting that self-generated action might not always lead to better recall than action observation (Feyereisen 2009). However, in Feyereisen (2009), both the self-generated action and action observation conditions led to better learning that reading verbal information alone. The specific effect of enactment on learning might also be dependent on the type of word that has to be learned. When learning the meaning of abstract verbs or locomotion verbs, enactment might not improve word learning (De Nooijer et al. in press).

Overall, the evidence from gesture observation, imitation, and enactment illustrates that building a link between previous motor experience and word meaning can play an important role in improving word learning. Out of the 17 studies reviewed that used gesture observation, imitation, or enactment, only one study found no difference in conditions where one would expect motor experience to aid learning (Feyereisen 2009; see Table 1). Critically, in this study, there was no difference between action observation and action imitation, but they did find a facilitative effect on memory for both action observation and action imitation compared to word reading alone. This still suggests an important role for motoric information. Similarly, one study did not find a benefit for action imitation for all words (De Nooijer et al. 2013), and two studies showed individual differences in the effectiveness of observing gesture compared to providing verbal information on word learning (De Nooijer et al. in press; Rowe et al. 2013). Still, no studies found that providing verbal information alone led to better word learning.

The motor experience of actually performing a meaning-related movement or gesture (i.e., imitation or enactment) appears to improve word learning. However, the effects of gesture observation and imitation may be sensitive to the type of target word (e.g., De Nooijer et al.

2013). Although individual differences are not always found with gesture-aided vocabulary techniques, when they are found, this could be due to the difficulty of the task or test. For example, results suggest that gestures are mostly beneficial in difficult tasks (McNeil et al. 2000), which suggests that gesture would be less helpful when the learning task is easy. But whether a task is difficult depends on what the learner finds difficult. Another possibility is that difficulty interacts with ability of the learner. For example, in some cases, action execution (i.e., imitation and enactment) appears to mainly benefit children with high verbal skills but not children with low verbal skills (De Nooijer et al. in press). This may be related to working memory resources needed to imitate or enact the movement. However, it is also possible that a type of cueing or segmenting of the information could put an end to this problem just like it has been shown to do with multimedia techniques (e.g., De Koning et al. 2009; Van Gog 2014; Spanjers, et al. 2010).

Summary of Vocabulary Techniques

In summary, there are many different techniques that are commonly used to help vocabulary development. The most important characteristics of improved vocabulary training appear to include (1) the amount of practice given to learn the words, (2) the richness of the information about the meaning of the words, (3) establishing ties between the students' own experience and knowledge and the words to be learned, and (4) the amount to which active processing was encouraged in the training (e.g., Beck, et al. 2002; Beck et al. 1987; Mezynski 1983; Nagy and Herman 1987; Stahl and Nagy 2006). The importance of the information about the meaning of the words, the establishment of ties between previous experience and the word to learn, and the amount to which active processing was encouraged in the training can all be explained (and enhanced) with an embodied view of vocabulary learning, as evidenced by the fact that out of 44 results summarized in this review, only three sets of results are not predicted by the embodied view (see Sweller, et al. 2011; Bull and Wittrock 1973; Acha 2009 summary line in Table 1; also note discussion of individual differences when using gesture observation and imitation above).

Generally speaking, the results suggest that each of these three factors contributes to making a congruent link between the words to learn and one's own experiences via the sensorimotor system.

Discussion, Questions, and Future Research

The results reviewed here suggest that the opportunity for a learner to make a congruent link from to-be-learned words to their own experiences via perceptual and motoric information is a key factor in effective vocabulary development. This link can occur via sensory information (e.g., keyword method, concept wheels, semantic word mapping, and illustrations) but importantly can also occur via action experiences (e.g., dynamic action or gesture observation, enactment, gesture, action imitation).

Enriching the Perceptual and Motoric Link in Vocabulary Techniques

Although not necessarily specifically designed with sensory information in mind, the more effective methods of vocabulary instruction techniques reviewed above typically provide a link

Table 1 (continued)

Authors	Key training comparison	Overall finding	Supports embodied cognition framework
Smeets and Bus (2014) Hald et al. (2014)	Dynamic illustration versus static pictures	Dynamic illustration was better	Yes
Kontra et al. (in press)	Physically interacting with objects versus observing objects in correct motion	Physically interacting with object improved learning of science concepts best	Yes
Tellier (2005) Goodwyn et al. (2000) De Nooijer et al. (in press) Rowe et al. (2013)	Observing gesture with verbal information versus verbal information only	Observing gesture improved learning De Nooijer: individual differences for all but one verb type Rowe: benefit of observing gestures mainly seen for low verbal ability children	Yes
Lindstromberg and Boers (2005) Allen (1995) Dijkstra and Kaschak (2006) Engelkamp et al. (1994) Casale (1985)	Enactment or pantomime compared to repeating verbal information	Enactment or pantomime led to best learning	Yes
Tellier (2008) Macedonia, et al. (2011) Kelly, et al. (2009) Macedonia (2003) Frick-Horbury (2002) De Nooijer et al. (2013)	Observing and imitating gesture versus only observing gesture	Observing and imitating better than observing only De Nooijer: possibly not for all word types	Yes
Feyereisen (2009)	Observing and imitating gesture, observing gesture, or verbal information only	No difference between observing and imitating versus just observing, but both were better than verbal information	Yes, partially

between sensory information and word meaning. For example, in the keyword method, the learner creates an interactive image that they can connect to the definition, thereby making a link to the visual sensory domain and possibly other sensory domains depending upon the exact image. To remember that *pont* in French means *bridge*, a learner could use the keyword method to link the meaning to both visual and olfactory information. For instance, if the learner thinks about the fact that *pont* is close to *point* and then creates an image where they are standing at the top point of a bridge (assuming a somewhat rounded bridge), they could imagine smelling the sea air on that bridge.

A comparable argument can be made with imagery training aimed at learning metaphorical meanings (e.g., Lindstromberg and Boers 2005). Furthermore, imagery training of action words has been shown to activate the premotor cortex (see Willems, et al. 2010), suggesting that imagery training may connect the meaning to motor information as well as sensory information (e.g., Lotze et al. 2003).

According to Dual Coding Theory (Sadoski 2005, 2009), techniques like this are effective because they allow learners to establish ties between the learners' own experience and the meaning via the activation of the verbal and nonverbal code. We believe this is a good start to describing what is needed to improve vocabulary training. However, we believe that this link can be enriched by more explicitly considering previous perceptual and, especially, motoric experience. Research reviewed here using movement observation, gesture observation, and imitation or enactment suggests that including motoric information might be the key (see Table 1 below for overview of findings supporting this conclusion).

Arguably, using action observation or imitation techniques in combination with more traditional (perceptual based) techniques could lead to greater vocabulary improvement. If meaning is based on the collection of bodily states (perceptions and actions) we have experienced in the past (e.g., Barsalou 1999), it follows that if you learn a new word by connecting it with multiple bodily states (perceptual and action), you will have a richer understanding of the word.

Within this framework, there is not a straightforward prediction about the limits of this possibility for learning. For example, as reviewed above, when people read words whose meaning are more related to either a specific perceptual modality (words about green objects) or action with a specific body part (kick, lick, pick), brain activation specifically occurs in an overlapping area to what one would use to actually perceive (e.g., see a frog versus pet a cat, Goldberg et al. 2006) or act out the denoted action (e.g., kick a ball versus lick ice cream; Tettamanti et al. 2005). It is likely that when learning a new word, information across different modalities including motoric information may well all contribute independently to development of a richer representation, while simultaneously preventing cognitive load problems that may occur when multiple pieces of information about word meaning (e.g., text and pictures) are presented in the same (usually visual) modality (e.g., Sweller et al. 2011). However, there are likely differences in how processes from different modalities would interact in vocabulary training depending upon characteristics of the learner and the words to be learned (e.g., Homer et al. 2008). For example, it has already been shown that people can differ in their preference for visual versus verbal learning (e.g., Mayer and Massa 2003).

The possibility of improving vocabulary techniques by more directly linking motoric experience and additional perceptual experience brings up several issues that would need to be explored. In particular, the importance of cognitive load, individual differences among learners, and differences in learning materials (e.g., word types) needs to be considered. We briefly consider each of these issues in turn, but specific empirical research on how these factors relate to ideas from embodied cognition are lacking.

Cognitive Load and Individual Differences

One concern of implementing many of the embodied cognition techniques for word learning is the individual differences that are sometimes found. For example, the finding that word meanings paired with pictures or iconic gestures appears to mainly help children with low verbal abilities (e.g., Rowe et al. 2013), but at the same time imitation and enactment may only benefit children with high verbal skills (e.g., De Nooijer et al. in press). We believe that this is not a cause to dismiss these techniques outright. The individual learner characteristics are important regarding the effectiveness of the strategies. Nonetheless, the overall findings using embodied cognition techniques such as multimedia, dynamic multimedia, and gestures hold for most learners. Furthermore, when individual difference has been found, it is typically not the case that the strategies did not work. It is more that different multimedia/gesture strategies worked differently for different learners (see Table 1 for a summary). Research on cognitive load and split-attention has already started to address how to solve this problem. Additional research is needed to better understand how different learners can benefit most from embodied cognition techniques to vocabulary development.

As mentioned above, it is still an open question whether the addition of information in other modalities (e.g., adding pictorial and/or motoric information to a verbal definition) has beneficial effects on learning because this leads to a richer mental representation of the learning material or because working memory load is reduced. Note that these two explanations are not mutually exclusive and that both factors may contribute to the beneficial effects on learning. Also note that adding information may actually increase cognitive load when it is not appropriately designed (e.g., induces split-attention) and that the amount of cognitive load imposed by the learning materials also depends on individual differences among learners, such as their prior knowledge, working memory capacity, or cognitive ability.

It may be the case that adding information in other modalities is more useful for novices (e.g., McNeil et al. 2000). Similarly, there may be differences in which combinations of perceptual information or motoric information lead to greater learning and less cognitive load than others. For example, it may be the case that adding olfactory information, action information, and visual information does not lead to greater cognitive load for some words. For example, this may be the case for words where olfactory information is important in the meaning (e.g., sea air). At the same time, the same combination of olfactory, action, and visual information may be less informative and lead to greater cognitive load for other words (e.g., computer screen). Related to this, there may be differential patterns of individual differences across the different modalities. For example, while gestures might be more useful for children with high verbal skills (e.g., De Nooijer et al. in press), other forms of action observation or other combinations of sensory information might be more useful for children with lower verbal skills.

This requires attention in future research, because at this point in time, embodied cognition theories are not explicit enough to say under what circumstances combining perceptual information (from more than one sensory modality) and motoric information would be beneficial for learning or would hamper learning. Furthermore, there has not been enough systematic vocabulary research looking specifically at multisensory and action-related word learning beyond just the visual modalities.

Effect for Different Word Types

Another interesting avenue for research is whether perceptual and motoric pieces of information are differentially effective for learning depending upon the type of words to be learned. For example, as discussed above, the effects of gesture observation, imitation, and enactment seem different for learning concrete versus abstract verbs in the first language (De Nooijer et al. 2013, in press). There may also be differences among nouns, verbs, adjectives, and prepositions, as well as idioms and metaphors. It seems intuitive that concrete nouns and verbs will be easier to relate to sensory and motor information than other words (*jump* compared to *justice*; see also Sadoski 2005 for a related argument). However, this relationship is not as straightforward as it may seem. Several recent studies have found evidence for improved learning of abstract words in a second language with a multisensory multimedia technique (Tsou et al. 2002, 2006). Note though that L2 learning is very different from L1 learning in the sense that the meaning of the word is already known but needs to be coupled to a new word. Furthermore, other evidence suggests that emotion plays a critical role in learning abstract words (Kousta et al. 2011). This suggests that using embodied techniques specifically related to simulating emotional states would be more effective for learning abstract words (see "Additional Question: the Role of Emotion" section below for more details of this argument). Similarly, evidence suggests that learning prepositions such as *to* and *for* can be improved by applying ideas from embodied cognition (Tyler et al. 2011). The same is true for learning metaphorical meanings (e.g., Boers 2013; Lindstromberg and Boers 2005). Combining this evidence suggests that incorporating ideas from embodied cognition could improve vocabulary training for more than just concrete nouns and verbs. However, to date, this positive effect of perceptual and motoric information on word learning is not consistently found across different word types, and future research should continue to make systematic comparisons in both L1 and L2 word learning. One possibility with regard to the differences found is that it may matter what the relationship is between the type of embodied technique applied and the meaning of the word. For example, it may be that manipulation knowledge, such as "knowing-how" to grasp and use a hammer, and functional knowledge, such as "knowing-that" a hammer is used for hammering, are linked to perceptual and motoric information differently (see Van Elk et al. 2014 for an extensive discussion of this difference and how it might matter for learning). Rather than word class being the critical issue in effective embodied learning strategies, it is whether the type of learning strategy matches well with the type of knowledge that is being taught.

Additional Effects to Consider

In the current review, we have considered different types of word learning techniques and how embodied cognition provides a compelling explanatory framework to explain the effectiveness of these techniques as well as some speculation of how actively incorporating motor and perceptual information may even improve some word learning methods (e.g., imagery). Nonetheless, there are many other factors which are typically not considered to be embodied that can also have an impact on learning, for example transfer-appropriate processing, testing effects, or distributed practice. Transfer-appropriate processing is the idea that we can remember things best when we retrieve the information under circumstances that are identical to the original learning experience (Morris et al. 1977). In neural terms, this has been described as a

function of the amount to which the same neural processes transfer correctly from the study experience to the memory test (see Schendan and Kutas 2007). The testing effect is the finding that the time on studying is more effective when some of this time is devoted to retrieving your knowledge (testing yourself) rather than just studying (Carrier and Pashler 1992). Distributed practice is the use of multiple study opportunities over a space of time. Using distributed practice has been shown to lead to very robust memory advantages (see Carpenter et al. 2012 for a review).

Although none of these effects are typically described within the embodied framework, they all fit with the framework and they can be combined with the embodied vocabulary techniques discussed in this review. For example, when planning to use gesture imitation as a means of vocabulary training, the learners should have the opportunity to perform the gesture at test, thereby keeping the circumstances that they learn and recall the information as constant as possible. We earlier discussed evidence suggesting a benefit for learning science concepts when the same neural processes (perceptual motor in this case) during learning occurred during recall (Kontra et al. in press). This is a good argument in favor of considering the environment of recall compared to the environment the information was learned, especially for these embodied techniques since the learning environment is often not just on paper, but test environments often are. Similarly, testing effects and distributed practice can be combined well with embodied techniques. For instance, if imagery is used to aid vocabulary learning, one could ask the learners to create different images for the same word over different practice sessions. Likewise, to apply testing effects, it would be better to devote some of the practice time to allow for learners to try to create images for target words when they do not have the definition in front of them. A better understanding of how effects like transfer-appropriate processing, testing effects, or distributed practice interact with embodied techniques of word learning is an interesting area for future research. At this point, we can only speculate, but it could be the case that applying some of these other findings to embodied learning techniques would greatly improve the effectiveness of the embodied techniques and possibly reduce individual differences by providing more support for the poorest learners.

Additional Question: the Role of Emotion

Before concluding, it should be noted that embodied cognition theories suggest that we understand our world by performing simulations based on previous perceptual, action, and *emotional* experiences. In this review, there has been no discussion of emotion. Although there is an extensive literature on the relationship between emotion and memory (see Kensinger and Schacter 2008), emotion, and embodied cognition (see Wilson-Mendenhall et al. 2013), only more recently is research emerging about the connection between emotion, embodied cognition, and language (see for example Glenberg et al. 2009; Havas et al. 2010). The relationship between providing related emotional information during vocabulary training is an area where very little is known. However, there is good reason to believe that emotion may play a critical role in learning abstract words. It has been argued recently that emotion plays a critical role in the representation of abstract concepts (Kousta, et al. 2011). In three experiments and a large-scale regression analysis of previous results, the authors demonstrate that the representation of abstract and concrete words includes sensory and motor experiential information, but abstract words are more emotionally valenced than concrete words. In particular, they argue that emotion plays a critical role for grounding abstract lexical concepts and that emotion provides

a bootstrapping mechanism for learning abstract lexical concepts (Kousta et al. 2011). Although the authors discuss this possibility using evidence from infant language acquisition, the same could hold for vocabulary development in school and later in life.

Conclusion

The goal of this review was to promote discussion about how current vocabulary training methods could be optimized by considering the evidence from research on (neuro)cognitive and educational psychology on word learning. In particular, we considered how mechanisms proposed by embodied cognition theories might provide a good basis for future vocabulary training experiments and methods. The idea that word learning can be improved by creating links from the to-be-learned words to the learners' own experiences and knowledge is not a new one (e.g., Nagy and Herman 1987). Although the evidence reviewed here suggests that vocabulary training methods could be optimized by incorporating a larger range of perceptual information, the more noteworthy innovative point is the importance of considering motoric information, as this has typically been ignored in standard vocabulary methods. To conclude, there is some evidence that learners can develop a better, deeper understanding of the meaning of words when a congruent link is made between the words to learn and ones' own motoric as well as perceptual experiences. This seems a fruitful area for future research, in which the boundary conditions of beneficial effects in terms of learner or material characteristics should also be taken into account.

References

Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology, 40*(1), 23–31. doi:10.1111/j.1467-8535.2007.00800.x.

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.

Anderson, R. C., & Kulhavy, R. W. (1972). Imagery and prose learning. *Journal of Educational Psychology, 63*, 242–243. doi:10.1037/h0032638.

Asher, J. J., & Price, B. S. (1967). The learning strategy of the total physical response: some age differences. *Child Development, 38*, 1219–1227. doi:10.2307/1127119.

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

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206–226). New York: Cambridge University Press.

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.

Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Erlbaum.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life*. New York: The Guilford Press.

Boers, F. (2013). Cognitive linguistics approaches to teaching vocabulary: assessment and integration. *Language Teaching, 46*, 208–224. doi:10.1017/S0261444811000450.

Boers, F. J., Piquer, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research, 13*, 367–382. doi:10.1177/1362168809341505.

Bull, B., & Wittrock, M. (1973). Imagery in the learning of verbal definitions. *British Journal of Educational Psychology, 43*, 289–293. doi:10.1111/j.2044-8279.1973.tb02269.x.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: review of recent research and implications for instruction. *Educational Psychology Review, 24*, 369–378. doi:10.1007/s10648-012-9205-z.

Carr, E. M., & Mazur-Stewart, M. (1988). The effects of the vocabulary overview guide on vocabulary comprehension and retention. *Journal of Reading Behavior, 20*, 43–62. doi:10.1080/10862968809547623.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory and Cognition, 20,* 632-642. Retrieved from http://laplab.ucsd.edu/publications/20

Casale, U. P. (1985). Motor imaging: a reading-vocabulary strategy. *Journal of Reading, 28*, 619-621. Retrieved from http://preview.tinyurl.com/nhq498n.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. doi:10.1017/S0140525X12000477.

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.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: guidelines for research and design. *Educational Psychology Review, 21*, 113–140. doi:10.1007/s10648-009-9098-7.

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. (in press). Words in action: using gestures to improve verb learning in primary school children. Gesture.

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.

Ekman, P. & Friesen, W. V. (1969). The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica, 1*, 49–98. Retrieved from http://www.paulekman.com/journal-articles.

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

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

Eppler, M. A. (1995). Development of manipulatory skills and the deployment of attention. *Infant Behavior and Development, 18*(4), 391–405.

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

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

Frick-Horbury, D. (2002). The use of hand gestures as self-generated cues for recall of verbally associated targets. *The American Journal of Psychology, 115*, 1–20. doi:10.2307/1423671.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 364*(1521), 1211–1221. doi:10.1098/rstb.2008.0300.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society London B Biological Sciences, 362*, 659–669. doi:10.1098/rstb.2006.2002.

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

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatorsensory voxels in all tested subjects: single-subject analysis of unsmoothed fMRI data. *Cerebral Cortex, 19*(6), 1239–1255. doi:10.1093/cercor/bhn181.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596. doi:10.1002/wcs.55.

Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education, 4*, 5–18.

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

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

Glenberg, A. M., Webster, B. J., Mouilso, E., & Lindeman, L. M. (2009). Gender, emotion, and the embodiment of language comprehension. *Emotion Review, 1*, 151–161. doi:10.1177/1754073908100440.

Glenberg, A. M., Goldberg, A. B., & Zhu, X. (2011). Improving early reading comprehension using embodied CAI. *Instructional Science, 39*, 27–39. doi:10.1007/s11251-009-9096-7.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience, 26*(18), 4917–4921. doi:10.1523/JNEUROSCI.5389-05.2006.

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.

Hald, L., Van den Hurk, M. & Bekkering, H. (2015). *Learning verbs more effectively through meaning congruent action animations. Learning and Instruction, 39,* 107-122. doi:10.1016/j.learninstruc.2015.05.010.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301–307.

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*(7), 895–900. doi:10.1177/0956797610374742.

Hegarty, M., & Kriz, S. (2007). Effects of knowledge and spatial ability on learning from animation. In R. Lowe & W. Schnotz (Eds.), *Leaning from animation* (pp. 3–29). Cambridge: Cambridge University Press.

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.

Homer, B. D., Plass, J. L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior, 24*, 786–797. doi:10.1016/j.chb.2007.02.009.

Jenkins, J. R., & Dixon, R. (1983). Vocabulary learning. Contemporary educational psychology. *Contemporary Educational Psychology, 8*, 237–260. doi:10.1016/0361-476X(83)90016-4.

Johnson, D. D., & Pearson, P. D. (1978). *Teaching reading vocabulary.* New York: Holt Rinehart & Winston.

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.

Kensinger, E. A., & Schacter, D. L. (2008). In M. Lewis & J. M. Haviland-Jones (Eds.), *Memory and emotion* (3rd ed., pp. (pp. 610--617).

Kontra, C., Lyons, D., Fischer, S. & Beilock, S.L. (in press). Physical experience enhances science learning. Psychological Science.

Kousta, S., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General, 140*(1), 14–34. doi:10.1037/a0021446.

Koziol, L. F., Budding, D. E., & Chidekel, D. (2011). From movement to thought: executive function, embodied cognition, and the cerebellum. *The Cerebellum, 11*(2), 505–525.

Lindstromberg, S., & Boers, F. (2005). From movement to metaphor with manner-of-movement verbs. *Applied Linguistics, 26*, 241–261. doi:10.1093/applin/ami002.

Lotze, M., Scheler, G., Tan, H.-R. M., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *NeuroImage, 20*, 1817–1829. doi:10.1016/j.neuroimage.2003.07.018.

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., 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 (German: Voice Movement Icons. Sensomotorische Encodierungsstrategie zur Steigerung der quantitativen und qualitativen Lerneffizienz bei Fremdsprachen) PhD Thesis. Salzburg: University of Salzburg.

Macedonia, M., & von Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biolinguistics, 6*, 393-416. Retrieved from http://www.biolinguistics.eu/index.php/biolinguistics/article/view/248/269.

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*(6), 982–998. doi:10.1002/hbm.21084.

Manzo, A. V. (1982). Subjective approach to vocabulary acquisition. *Reading Psychology, 3*(2), 155–160. doi:10.1080/027027182003020.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25–45. doi:10.1146/annurev.psych.57.102904.190143.

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833. doi:10.1037/0022-0663.95.4.833.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 84*, 389–460. doi:10.1037/0022-0663.86.3.389.

McGregor, K. K., Sheng, L., & Ball, T. (2007). Complexities of word learning over time. *Language, Speech, and Hearing Services in Schools, 38*, 1–12. 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.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: effects of vocabulary training on reading comprehension. *Review of Educational Research, 53*, 253–279. doi:10.3102/00346543053002253.

Miyashita, Y. (1995). How the brain creates imagery: projection to primary visual cortex. *Science, 268*(5218), 1719–1720. doi:10.1126/science.7792596.

Montessori, M. (1966). *The human tendencies and Montessori education*. Amsterdam: Association Montessori Internationale.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*, 519–533. doi:10.1016/S0022-5371(77)80016-9.

Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. In M. McKeown & M. Curtiss (Eds.), *The nature of vocabulary acquisition* (pp. pp. 19–pp. 35). Hillsdale: Erbaum.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. Washington, DC: U.S. Government Printing Office.

Needham, A. (2000). Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognition and Development, 1*(2), 131–156. doi:10.1207/S15327647JCD010201.

Nilsson, L. (2000). Remembering actions and words. In T. Endel, I. Fergus, M. Craik, & M. Craik (Eds.), *The Oxford handbook of memory* (pp. pp. 137–pp. 148). New York: Oxford University Press; US.

Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2012). Interplay between action and movement intentions during social interaction. *Psychological Science, 23*, 30–35. doi:10.1177/0956797611424163.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Paivio, A. (1986). *Mental representations: a dual coding approach*. Oxford, England: Oxford University Press.

Paivio, A. (2007). *Mind and its evolution: a dual coding theoretical approach*. New York: Psychology Press.

Pearson, D., Hiebert, E., & Kamil, M. (2007). Vocabulary assessment: what we know and what we need to learn. *Reading Research Quarterly, 42*, 282–296. doi:10.1598/RRQ.42.2.4.

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2012). Computational grounded cognition: a new alliance between grounded cognition and computational modeling. *Frontiers in Psychology, 3*, 612. doi:10.3389/fpsyg.2012.00612.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*, 25–36. doi:10.1037/0022-0663.90.1.25.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior, 19*, 221–243. doi:10.1016/S0747-5632(02)00015-8.

Pressley, M. G. (1976). Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology, 68*, 355–359. doi:10.1037/0022-0663.68.3.355.

Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: mnemonic and contextual approaches. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 107–127). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Proverbio, A. M., Riva, F., & Zani, A. (2009). Observation of static pictures of dynamic actions enhances the activity of movement-related brain areas. *PLoS-ONE, 4*(5), e5389. doi:10.1371/journal.pone.0005389.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*, 576–582. doi:10.1038/nrn1706.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*(5), 351–360. doi:10.1038/nrn2811.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & llmoniemi, 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.

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

Rowe, M., Silverman, R., & Mullen, B. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a new language. *Contemporary Educational Psychology, 38*(2), 109–117. doi:10.1016/j.cedpsych.2012.12.001.

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

Sadoski, M. (2009). Embodied cognition, discourse, and dual coding theory: new directions. In J. Renkema (Ed.), *Discourse, of course: an overview of research in discourse studies* (pp. 187–196). Amsterdam: John Benjamins Publishing Co.

Sadoski, M., & Paivio, A. (1994). A dual coding view of imagery and verbal processes in reading comprehension. In R. Ruddell, M. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (pp. 582–601). Newark, DE: International Reading Association.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: a dual coding theory of reading and writing*. Mahwah, NJ: Lawrence Erlbaum.

Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. In R. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 1329–1362). Newark, DE: International Reading Association.

Sadoski, M., Paivio, A., & Goetz, E. (1991). A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly, 26*, 463–484. doi:10.2307/747898.

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

Schendan, H. E., & Ganis, G. (2012). Electrophysiological potentials reveal cortical mechanisms for mental imagery, mental simulation, and grounded (embodied) cognition. *Frontiers in Psychology, 3*, 329. doi:10.3389/fpsyg.2012.00329.

Schendan, H. E., & Kutas, M. (2007). Neurophysiological evidence for transfer appropriate processing of memory: processing versus feature similarity. *Psychonomic Bulletin & Review, 14*(4), 612–619. doi:10.3758/BF03196810.

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.

Schwartz, R. M. & Raphael, T. E. (1985). Concept of definition: a key to improving students' vocabulary. *The reading teacher, 39*, 198-205. Retrieved from http://www.jstor.org/stable/20199044.

Silverman, R. (2013). Investigating video as a means to promote vocabulary for at-risk children. *Contemporary Educational Psychology, 38*(3), 170–179. doi:10.1016/j.cedpsych.2013.03.001.

Smeets, D. J. H., & Bus, A. G. (2014). The interactive animated e-book as a word learning device for kindergartners. *Applied Psycholinguistics*. doi:10.1017/S0142716413000556.

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.

Spanjers, I. E. A., Van Gog, T., & Van Merriënboer, J. J. G. (2010). A theoretical analysis of how segmentation of dynamic visualizations enhances students' learning. *Educational Psychology Review, 22*, 411–423. doi:10.1007/s10648-010-9135-6.

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior, 15*, 33–50. doi:10.1080/10862968309547495.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: a model-based meta-analysis. *Education & Educational Research, 56*, 72–110. doi:10.3102/00346543056001072.

Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings. The literacy teaching series*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Szczepaniak, R., & Lew, R. (2011). The role of imagery in dictionaries of idioms. *Applied Linguistics, 32*, 323–347. doi:10.1093/applin/amr001.

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: DIDCOG2005 [Document of the first conference of cognitive didactics]* (proceedings on CD-Rom). Toulouse, France: Université de Toulouse–Le Mirail.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture, 8*, 219–235. doi:10.1075/gest.8.2.06tel.

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

Tsou, W., Wang, W., & Li, H. (2002). How computers facilitate English foreign language learners acquire English. *Computers & Education, 39*(4), 415–428. doi:10.1016/S0360-1315(02)00078-7.

Tsou, W., Wang, W., & Tzeng, Y. (2006). Applying a multimedia storytelling website in foreign language learning. *Computers & Education, 47*, 17–28. doi:10.1016/j.compedu.2004.08.013.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies, 57*, 247–262. doi:10.1006/ijhc.2002.1017.

Tyler, A. Mueller. C. & Ho, V. (2011) Applying cognitive linguistics to learning the semantics of English prepositions *to, for*, and *at*: an experimental investigation. VIAL, *Vigo International Journal of Applied Linguistics, 8,* 181-205. Retrieved from http://webs.uvigo.es/vialjournal/number_8.html.

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

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

Van Elk, M., Van Schie, H., & Bekkering, H. (2014). Action semantics: a unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of Life Reviews, 11*, 220–250. doi:10.1016/j.plrev.2013.11.005.

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263–278). New York: Cambridge University Press.

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.

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

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. doi:10.3758/BF03196322.

Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (2013). Situating emotional experience. *Frontiers in Human Neuroscience, 7*(164), 1–16. doi:10.3389/fnhum.2013.00764.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy, 13*(3), 229–248. doi:10.1080/15250000802004437.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*, 244–262. doi:10.1016/j.cognition.2012.06.016.

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.