Learning and Instruction 39 (2015) 107–122

Contents lists available at ScienceDirect

Learning and Instruction

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

Learning verbs more effectively through meaning congruent action animations

Lea A. Hald [a, *], Marianne van den Hurk [b], Harold Bekkering [a]

[a] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[b] Pedagogical Sciences, Education Science, Radboud University Nijmegen, Postbus 9104, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
26 May 2015
Accepted 30 May 2015
Available online 16 June 2015

Keywords:
Word learning
Action observation
Motor simulation
Embodied cognition
Action-Based Language theory

ABSTRACT

The current study investigates the effectiveness of learning words while displaying meaning congruent animations. We explore whether learning words with animation is sensitive to properties known to influence action understanding. We apply an embodied cognition framework and predictions from a recent theory about language and action (Action-Based Language theory, Glenberg & Gallese, 2012). The current study aims to investigate whether dynamic animations add to word learning (Experiment 1) and what the linguistic relation between the dynamic animation and the word learning is (Experiment 2). Results indicate that meaning congruent animations improved verb learning compared to meaning incongruent animations when measured by a recognition task. When measured by an active recall task, congruent animations led to better learning than static pictures. In both measures, meaning congruent animations support word learning. Experiment 2 replicates and extends this and suggests that highlighting conceptual information related to the dynamic action (such as the goal) improves word learning further. The findings are in line with Action-Based Language theory, which suggests that children are able to make better simulations of an action during learning when supported by meaning congruent animations. Highlighting conceptual information additionally supports this learning process.

Language education is currently enriched by multimodal instructions to improve the learning rate of new words. Although employing multimedia aspects like pictures has been shown to be more effective for word learning than text alone for most learners (Smith, Stahl, & Neil, 1987; Smith, Miller, Grossman, & Valeri-Gold, 1994; Rowe, Silverman, & Mullan, 2013; see Stahl & Nagy, 2006; for a review) not much is known about the underlying neurocognitive mechanisms that improve learning. The current study aims to investigate whether dynamic animations add to word learning (Experiment 1) and, what the linguistic relation between the dynamic animation and the word learning is (Experiment 2). The theoretical starting point of these investigations is the Action-Based Language theory (Glenberg & Gallese, 2012) which suggests that perceiving congruent motor actions is important for word learning in general.

1. Animations and word learning

Much of the research using dynamic multimedia for word learning involves gesture, which has to be imitated by the participant, or the participant has to do some other type of movement. An exception to this is seminal work by Plass, Chun, Mayer, and Leutner (1998, 2003). They used multimedia (video recordings and pictures) to explore aspects of individual differences related to learning and cognitive load. Using an interactive multimedia program, American college students learning German as a second language were presented with a short story in German. For a selection of key words in the story, students could access an annotation about the meaning of that word. The annotations were either verbal or visual, where the visual condition contained either pictures or short video recordings. In Plass et al. (1998) the students choose which type of instruction they received while in Plass et al. (2003) the annotation mode was predetermined between subjects. Overall, there were improvements in recall of individual vocabulary items when students used both visual and verbal annotations (Plass et al., 1998) but this was dependent on individual differences in spatial and verbal abilities (Plass et al., 2003). However, most

* Corresponding author. Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands.
E-mail addresses: l.hald@donders.ru.nl, lea.hald@gmail.com (L.A. Hald), m.vandenhurk@pwo.ru.nl (M. van den Hurk), h.bekkering@donders.ru.nl (H. Bekkering).

http://dx.doi.org/10.1016/j.learninstruc.2015.05.010
0959-4752/

important for the current study, Plass et al. did not report on the difference in learning between still pictures and short video recordings. Video recordings and pictures were grouped together in the visual annotations.

Results from other research using videos or animations for vocabulary learning have been mixed. For example, Neuman and Koskinen (1992) explored the effectiveness of watching television with or without subtitles in comparison to reading and listening to the same information or simply listening. They found the greatest gains in learning when participants watched subtitled television, suggesting that animations can aid word learning when presented together with the word. More recent research looking at vocabulary learning from watching educational television showed similar gains in learning words (Linebarger, Moses, Liebeskind & McMenamin, 2013). However, the effectiveness of subtitles (displaying the word to learn) varied based on the child's socioeconomic status: Low socioeconomic status children were not helped by the subtitles. Middle-class socioeconomic status children were helped by the subtitles, but the subtitles did not lead to additional gains when there was repeated exposure (Linebarger et al., 2013).

Research by Baltova (1994) used either combined video-audio clips or audio-only clips to teach French to English speaking 8th-grade students and found no significant difference. Furthermore, Sun and Dong (2004) found animation in itself was ineffective for teaching English to Chinese children. Only when the animations were paired with textual support did they find improvements in learning. Other studies using video and subtitles have shown that subtitles in a second language can aid perceptual learning (Mitterer & McQueen, 2009).

In summary, meaning congruent video or animations may help with word learning, but it is unclear whether animations can provide additional benefits compared to pictures. It should be noted that none of the research above was intended to explore the added benefit that meaning congruent animations may bring over meaning congruent pictures, but this researched focused mainly on visual versus verbal information instead (e.g., Plass et al., 1998).

There is a long tradition of exploring the use of animations versus pictures for learning other types of information, such as procedural motor knowledge, declarative knowledge or problem solving (see Höffler & Leutner, 2007 for a review). In their meta-analysis Höffler & Leutner (2007) conclude that representational animations improve learning more than decorative animations. Furthermore they suggest that learning procedural motor knowledge benefits the greatest from animation, followed by learning declarative knowledge and then problem solving. The largest improvement with learning procedural motor knowledge with animations could be due to the fact that these animations illustrated motor actions. This would be in line with our hypothesis about word learning via motor actions. There is some recent support for this possibility. For example, Brucker, Ehlis, Häußinger, Fallgatter and Gerjets (2015) found that university students could learn patterns of fish movements' best when they watched animations of hands gesturing the correct movements. Similarly, it has been suggested that learning biological or manipulative tasks may be better learned by animations, where learning non-biological or symbolic tasks may be better learned by static pictures (Castro-Alonso, Ayres & Paas, 2014). This suggestion makes it particularly interesting to investigate where learning word meanings fit in this continuum.

None of the research on using animations for word learning systematically used animations that demonstrated motor actions. However, there is an independent reason to believe that watching motor actions may be useful for word learning. The embodied cognition framework and in particular, a new theory specifically outlining the coupling of language and action (Glenberg & Gallese, 2012), suggests that perceiving congruent motor actions is important for word learning.

1.1. Motor actions and embodied cognition

Why might motor involvement in general help with word learning? One explanation is given by the theory of embodied cognition: The basic idea is that our knowledge about objects and other referents in the world, and the way in which we understand the world, is grounded in our perception and actions, rather than some type of abstract symbols (Barsalou, 1999). For example, our concept of *bird* is based on the collection of bodily states (perceptions, actions and emotions) we have experienced with birds. Broadly speaking, these previous perceptions are partially reactivated (simulated) in order to understand *bird*. 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; Taylor & Zwaan, 2009; Glenberg & Kaschak, 2002). While it provides us with an interesting starting point, the framework of embodied cognition does not necessarily lead to any specific predictions about the use of meaning congruent animations compared to pictures for word learning.

1.2. Action-Based Language theory

Recently a formal theory has been proposed to account for the relationship between the action aspect of embodied cognition and language, the Action-Based Language theory (ABL, Glenberg & Gallese, 2012). The ABL theory includes a mechanism of motor control and paired controller/predictor models to account for language learning, comprehension and production. The ABL theory makes use of two well-established models of motor control as the framework for the relationship between action and language, the MOSAIC and HMOSAIC theories (Haruno, Wolpert, & Kawato, 2003). These models propose a controller and a predictor. The controller computes motor commands from a representation of goals and context, while a predictor predicts the motor and sensory consequences of actions and sends error information to the controllers to update motor commands as needed.

What is novel about the ABL theory is that Glenberg and Gallese propose that the brain solves the problem of how to control contextually-appropriate action and contextually-appropriate language via the same process. They propose that the opportunistic sharing of action control and language processing is possible through mirror neurons. The theory links verb meaning directly to the motor actions associated with the word via the controllers. In addition, meaning is linked directly to the predictions (from the predictors) that can be derived from simulating the action. The ABL theory predicts that meaning congruent animations would allow the child to encode the word through a simulation so that better predictions can be made. For example, if learning the word *chiseling*, a simulation would lead to predictions about how a chisel can be used, what would result from using a chisel, and what type of objects would afford chiseling. This type of information could be predicted much better when viewing an animation. Therefore, encoding the word assisted by animation should lead to a more comprehensive understanding of the word meaning.

In sum, according to the ABL theory, understanding what it means *to chisel* involves the sensory-motor feedback of chiseling. This predicts that performing the action yourself would provide the strongest environment for this learning to occur. But the difference