# NEURAL CORRELATES OF IMAGERY-BASED FOREIGN WORD LEARNING

A Thesis

Presented to

The Faculty of the Department

of Psychology

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

Kailyn A. L. Bradley

December, 2012

NEURAL CORRELATES OF IMAGERY-BASED FOREIGN WORD LEARNING

Kailyn A. L. Bradley

**APPROVED:**

Arturo E. Hernandez, Ph.D.
Committee Chair

Merrill Hiscock, Ph.D.

Hanako Yoshida, Ph.D.

John W. Roberts, Ph.D.
Dean, College of Liberal Arts and Social Sciences
Department of English

NEURAL CORRELATES OF IMAGERY-BASED FOREIGN WORD LEARNING

---

An Abstract of a Thesis

Presented to

The Faculty of the Department

of Psychology

University of Houston

---

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

---

By

Kailyn A. L. Bradley

December, 2012

**ABSTRACT**

The purpose of this study was to investigate whether an interactive imagery-based learning paradigm would facilitate foreign word learning more than picture imagery. Functional magnetic resonance imaging (fMRI) was used to examine the cognitive constructs underlying these pedagogies. Additionally, native language proficiency and imaging ability were evaluated as predictors of successful vocabulary learning. English monolinguals were trained on novel German vocabulary using two different imagery-based learning techniques for four consecutive days. Picture imagery training required participants to visualize the picture of an object presented with a novel word, whereas interactive imagery training required visualization of a first-person interaction with the object. After training, fMRI was used to assess the neural correlates of training during an auditory recognition task. Two weeks post-scan, vocabulary retention was assessed. Behavioral results revealed a training order effect in which undergoing picture imagery training first provided scaffolding that increased accuracy in interactive imagery training only during the first session. Additionally, English proficiency but not imaging ability predicted high vocabulary test scores. Moreover, imaging results revealed distinct neural patterns related to the different learning techniques. Words learned through picture imagery strongly activated basic visual processing regions, whereas activity for words learned through interactive imagery was greatly reduced and present in motor resonance and cognitive control regions. Taken together, these findings support the conclusion that it is initially more difficult for adults to adapt to using interactive imagery to learn vocabulary and that individual differences may predict successful learning in addition to training method.

**TABLE OF CONTENTS**

# LIST OF TABLES

## LIST OF FIGURES

**LIST OF APPENDICES**

## Introduction

How do we acquire a second language and why is it so difficult for adolescent and adult learners to achieve proficiency levels equivalent to young children? These questions hold significance for many sub-disciplines of psychology and educational research. While the literature shows that being immersed in another country and culture is the key to successful language acquisition (Sanz, 2005) it is not always convenient or possible to do. Therefore, there has been much interest from educators, policy makers, and language researchers to determine what combination of factors results in the greatest gain in second language skills without leaving the classroom. To date there have been no studies that evaluate the neural mechanisms underlying embodied learning methods that might aid second language vocabulary learning. Therefore, there is a gap in the literature that calls for examination of these different learning methods and the neural mechanisms that drive successful second language acquisition.

The question guiding this study has roots in the embodied theory of language acquisition. Is it enough to simply have multimodal input of a new word in order to create a concrete conceptual representation, or can perceiving interaction through mental imagery improve learning by creating a more embodied experience? This study will examine the cognitive constructs that underlie learning methods that use different types of mental imagery in the early stages of foreign language vocabulary acquisition. Specifically, this study examines if mental imagery can create a more embodied learning experience that might facilitate vocabulary learning beyond more traditionally used association methods that require learners to simply match pictures with words. In the following review of the literature, the embodied theory of language acquisition will be discussed, followed by studies that have linked motor actions and imagery to language representation. Lastly, the contributions of both pedagogy and individual differences in vocabulary learning will be addressed.

**Embodied Language Learning**

Embodied language theorists propose that word meanings are grounded in representations of perceptual, motor, and affective experiences (Barsalou, 1999, 2008; Semin & Smith, 2008; Glenberg & Kaschak, 2002; Bergen, Lau, Narayan, Stojanovic, & Wheeler, 2010; Glenberg, Goldberg, & Zhu, 2011). Embodied cognition grew out of perceptual symbol systems theory and the work of Barsalou (1999, 2008), who suggests that higher-order cognition, such as language comprehension, is the product of the construction of perceptual simulations. In this theory, representations of perceptions are extracted from experience, stored in memory, and manipulated to create understanding and meaning.

Building on the perceptual symbol systems theory, Glenberg and Robertson (1999) proposed the *indexical hypothesis*, which states that meaning is based not only on perception, but also on action. According to the *indexical hypothesis*, words are mapped to perceptual symbols (Glenberg & Kaschak, 2002) and because we are biomechanical beings, we infer meanings about objects by integrating affordances to accomplish action-based goals. Glenberg and colleagues (2011) propose that interacting with the environment can provide these particular embodied experiences that allow us to infer meaning; in other words, by using any of our senses to explore an environment, we can create memories of those experiences that are then easier to recall.

In addition to simply comprehending single words, comprehending sentences also involves mapping words to experiences and requires syntax to guide the integration of those experiences (Glenberg & Kaschak, 2002; Glenberg et al., 2011; Kaschak & Glenberg, 2000). In a series of experiments, children with reading comprehension problems were found to focus too narrowly on decoding individual words instead of mapping words to experiences for greater understanding of meaning (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004; Glenberg et al., 2011). Glenberg and colleagues found that manipulation of objects in a sentence on a computer screen or in children's minds greatly improved comprehension without actually having to manipulate actual objects. This research suggests that learning to map novel words to concepts may also be improved by interaction with objects in a virtual or imagined space, as opposed to having to interact with the physical world. These conclusions have led to the suppositions in this study, in which it is hypothesized that mentally manipulating objects may provide a similarly embodied experience to aid the creation of conceptual representations in foreign word learning.

While the interventions in Glenberg and colleagues' (2004; 2011) studies were meant to aid children's first language reading comprehension, they have implications for second language research as well. It is possible that second language learners also focus too narrowly on phonological and orthographic relationships in the beginning of vocabulary learning, rather than assigning meaning to words. It is often the case that second language learners acquire new vocabulary by mapping words in their native language (L1), which are already linked to conceptual representations, onto the new language (L2) (Hernandez, Li, & MacWhinney, 2005). According to this theory, as L2 forms gain strength, links can be made from L2 to the concept directly, without having to access the L1 representation. This mechanism of learning may not always be beneficial, as constant activation of the first language can create interference. Learning words by creating entirely new links between L2 and a concept, skipping the link to L1, is much closer to how infants and children learn a first language. If second language vocabulary learning can be made more similar to first language learning, by emphasizing the conceptual relationship between the new word and the object, skipping the link to L1, novel word learning and retention might be enhanced. Therefore, in this study, the L1 translation of novel words is never shown in order to de-emphasize this link. Participants are trained to make direct connections between novel words and objects using two types of imagery, one of which is possibly more embodied. In the next section, we discuss further evidence that language processing in general is embodied, through examination of the relationship between motor actions and language representation. If language learning is naturally an embodied process, then it is hypothesized that a more embodied learning method might be more successful for long-term retention of vocabulary information than a simpler picture imagery association method.

**Motor Actions and Language Representation**

Much of the current literature on embodied language comprehension focuses on the role of mirror neurons in the motor cortex in learning and processing action words (James & Maouene, 2009; Fischer & Zwaan, 2008). Fischer and Zwaan (2008) suggest that comprehending verbal descriptions of actions depends on internal simulation of the described action. For example, the recall and recognition of verbs compared to adjectives results in activation of mirror neurons in the motor cortex, a phenomenon known as motor resonance, which suggests the embodied nature of action word representation (James & Maouene, 2009; Fischer & Zwaan, 2008). These mirror neurons have been identified in the ventral premotor and inferior parietal brain regions of macaque monkeys both in response to performance of an object-directed action and after observance of the same type of object-directed action by another (Di Pellegrino, Fadiga, Fogassi, & Gallese, 2001). Fischer and Zwaan (2008) suggest that in humans, mirror neurons are less dependent on visually accessible action information and may be more broadly distributed in prefrontal language regions, specifically Broca's area, which is also the speech center in humans. In other words, not only observance of actual object-directed actions stimulates these cells, but so do sounds associated with actions (Keysers et al., 2003), passive observation of non object-directed gestures (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995), and static images of speech actions such as lip postures (Nishitani & Hari, 2002).

Several recent studies have examined motor resonance in first language processing and have shown that in adults, reading verbs activates the same regions of the motor cortex that are used during that actual action performance (Fischer & Zwaan, 2008; Pulvermuller, Harle, & Hummel, 2001; Pulvermuller, Hauk, Nikulin, & Ilmoniemi, 2005). This is also true in children (James and Maouene, 2009). Other lexical categories, including nouns and adjectives may also show motor resonance. Nouns that refer to objects that can be physically manipulated have been shown to activate the motor system (Saccuman, Cappa, Bates, Arevalo, Della Rosa, Danna, & Perani, 2006). As long as a word can be strongly associated with an action, similar motor resonance seems to occur (Saccuman et al., 2006; Arevalo, Perani, Cappa, Butler, Bates, & Dronkers, 2007).

From the above-described literature, it seems plausible that language processing may be embodied, and therefore, a more interactive learning experience might facilitate acquisition and retention of novel second language words by creating a stronger multimodal link between the word form and semantic representation. Many of the same motor related brain mechanisms employed during action production appear in many types of language processing. This literature suggests then that it is not necessary to actually produce an action to activate perceptual systems associated with an action. For instance, seeing someone else perform an action can activate these same systems. With this in mind, could it be possible to create an embodied experience that facilitates word learning using mental imagery? Several studies reviewed in the next section suggest this is the case.

**Imagery in Learning**

The idea of using imagery to enhance learning is not new. Literature examining the benefits of imagery in learning extends back several decades. Paivio (1971; 1986; 1991) first introduced the Dual-Coding Theory (DCT) that proposed the existence of two facets of cognition in learning, verbal associations and visual imagery. According to DCT research (Sadoski, 2005; Cohen and Johnson, 2011), paired-associate vocabulary learning, in which a word is paired with a visual representation, has proven successful because information that is stored in multiple codes (verbal and visual) is easier to retrieve than when it is simply encoded through the language system alone. Lutz and Lutz (1977) further explored factors that influenced paired-associate learning, specifically the effects of images depicting actions compared to static pictures. They found that an interactive image (i.e. a picture showing action) facilitated vocabulary learning above and beyond a simple static, non-interactive picture. This suggests that the hint of object-directed action can facilitate word learning.

More recent literature has begun to assess the usefulness of self-generated mental imagery in first language vocabulary learning and reading comprehension. Cohen and Johnson (2011) evaluated vocabulary learning through text only, text and picture pairs, or text and mental imagery pairs. While they did not find any significant differences in accuracy between the three methods across all vocabulary categories, in the most difficult science word category, mental imagery paired-associate learning was the most successful technique, suggesting that self-generated mental imagery can provide an additional code through which difficult vocabulary information can be stored and retrieved. Similarly, Glenberg, Gutierrez, Levin, Japuntich, and Kaschak (2004) found that both physical manipulation and imagined manipulation of objects in a sentence resulted in stronger memory and better comprehension of the material than simply rereading sentences. These results support the *indexical hypothesis* and suggest that mental imagery can create the necessary perceptual experience to promote embodied language learning.

Other non-language related studies have also shown mental imagery to be beneficial in learning. Schwartz and Black (1999) suggest that people make physical inferences about objects by acting on them, a principle of embodied learning. In a series of experiments, they had people judge physical properties of objects, for example, tilting glasses of different widths to see whether fluid would spill out at specified angles.

Physically manipulating objects allowed participants to make accurate judgments, but so did mentally imagining the action of tilting the containers. Additionally, mental imagery has successfully been used for decades to train athletes through imagined practice (Jones & Struth, 1997) and musicians through goal imaging of musical compositions (Woody, 2006). Mental imagery can therefore be used to rehearse information without having to actually perform the actions associated with it.

Given the above review of the imagery literature, it is evident that several studies support the use of picture imagery or mental imagery as a means of facilitating learning; however, the question still remains, can embodied instructional techniques translate to foreign language vocabulary learning in adults? Results of a recent study fail to fully address this question. Shen (2010) compared learning effects of two novel word encoding methods, one of which gave beginning adult Chinese learners verbal vocabulary information, and one that gave additional imagery, either pictures or video clips depicting a person physically acting out the actions associated with the words. Shen's (2010) results showed that for concrete words, added imagery did not have a superior effect on retention of word meanings; however, for more abstract words, added imagery facilitated retention of the meaning of words. These results give partial support to the idea that imagery can be used to create an embodied experience that may facilitate second language word learning. However, little distinction was made between how the different types of imagery facilitated learning, and the learner was not an active participant in either type of imagery learning, which is paramount to theories of embodied cognition. Therefore, further investigation in needed. We will add to the imagery literature by examining if interactive mental imagery (the mental manipulation of objects) can create a more embodied experience that might facilitate learning and retention of foreign language vocabulary information above a more traditional paired-associate picture imagery learning method.

**Individual Differences in L2 Learning**

While it is clear from the review of the imagery literature that training method might influence second language vocabulary acquisition, it is also possible that individual differences will contribute to variability in learning outcomes. In a comprehensive review of individual differences in language learning, Sparks and Ganschow (2001) suggests that over the last almost century, many factors have been identified as potential predictors of successful foreign language learning. Very early on, Symonds (1930) suggested that three factors influenced a student's ability to learn a new language: performance in the native language, general intelligence, and ability to learn quickly. Similarly, the military initially used education outcomes as predictors of successfully language learning (Spolsky, 1995). Over time, more comprehensive models have been formed to gauge what factors predict who might excel at language learning. Several of these factors include a learner's desire and motivation to learn, anxiety, ability to decode words and discriminate phonemes, ability to handle complex grammatical rules, enhanced rote memory, and native language skill (Sparks and Ganschow, 2001).

Currently, some of the most investigated individual differences in foreign language learning are native language skill and phonological decoding ability. The finding that native language ability (oral and written proficiency, word decoding, literacy and reading skills, etc.) is correlated with the ability to learn foreign languages has been replicated many times (Sparks and Ganschow, 1991, 2001; Sparks, Patton, Ganschow, Humbach, & Javorsky, 2006; Sparks, Patton, Ganschow, & Humbach, 2012; Sparks, 2012; Dufva & Voeten, 1999). Phonological decoding ability, often considered a component of native language skill, has also been shown to greatly influence language learning ability (Sparks, Ganschow, Patton, Artzer, Siebenhar, and Plageman, 1997; Sparks, Patton, Ganschow, and Humbach, 2009, 2011). Meschyan and Hernandez (2002) also found that native language decoding ability predicted adult second language competency. Together, these studies suggest that a natural awareness of phonological-orthographic relationships is highly correlated with the ability to decode words into sounds and successfully learn a foreign language. Therefore, in the current study, it is predicted that individuals' native language skill will be related to learning outcome. More specifically, those individuals with higher receptive and expressive English proficiency scores, as assessed by the Woodcock-Munoz Language Proficiency Battery, will be better vocabulary learners (higher vocabulary scores).

While much of the language learning literature supports the hypothesis that high native language proficiency will be correlated with novel vocabulary learning, there is very little research that investigates how imaging ability will affect vocabulary learning in paradigms that rely on different types of imagery. In an imagery study, Ernest and Paivio (1969) did not find a strong relationship between imaging ability and verbal skills. However, a follow-up study found otherwise. Ernest and Paivio (1971) recorded reaction times as individuals either imagined an image that went with a word or said another word associated with it. They found that although not quite significant, latency differences between good and poor imagers were greater for the imagery task than for the verbal association task. Additionally, in general, overall association time was related to imaging ability; individuals that were better imagers had shorter association latencies regardless of task. Moreover, Di Vesta and Ross (1971) also found that imagery ratings, both of word stimuli and as an individual difference factor significantly affect paired-associate learning outcomes. Therefore, in the current study, we hypothesize that individual differences in imaging ability as assessed by Marks' (1973) Vividness of Visual Imagery Questionnaire (VVIQ) will be related to vocabulary learning.

## Current Study Goals

Given the above review of the literature, this study seeks to examine several aspects of adult foreign language vocabulary learning. Specifically, will an interactive mental imagery paradigm create an embodied experience and facilitate vocabulary learning and retention beyond a simpler paired-associate picture imagery learning method? Using fMRI, the neural correlates of these learning methods will be examined to better understand the mechanisms behind learning. Additionally, this study will evaluate the contribution of individual differences to learning outcomes, expressly native language proficiency and imaging ability.

## Hypotheses

### Training method.

1. Both picture imagery and imagined interaction training will result in vocabulary learning, as current literature supports the successful use of imagery in training throughout several domains (vocabulary learning, reading comprehension, music, sports); however, interactive imagery training will result in higher

vocabulary scores and greater retention of vocabulary information after two weeks, as it is thought to be more embodied.

2. Motor resonance brain areas (i.e. inferior parietal, inferior frontal, motor and premotor cortices) will be more strongly activated in response to words learned through imagined interaction training. However, words learned through picture imagery will show activity in more basic visual processing areas and may also show activity in association brain regions such as the supramarginal gyrus (inferior parietal cortex), as this paradigm also involves the integration of multimodal (visual and verbal) information.

**Individual differences.**

3. First language (L1) proficiency will be correlated with foreign vocabulary learning. Learners with high L1 proficiency will learn more novel vocabulary words.
4. Imaging ability, as assessed through the Vividness of Visual Imagery Questionnaire (VVIQ), will be correlated with vocabulary learning performance. Better imagers (lower scores on the VVIQ) will show enhanced learning (higher vocabulary test scores) and retention of vocabulary in both paradigms, but more robustly for interactive imagery training.

## Experiment 1

The focus of this study was to determine the effects of different types of mental imagery training on foreign word learning in adults. In order to ensure ease of imageability of words in this study, Experiment 1 was conducted to gather ratings on nouns representing objects with high familiarity, manipulability, and sensory features, in addition to gathering reaction times for different types of imaging tasks to create the training paradigms used in Experiment 2. Nouns were chosen instead of verbs, as they are most often learned first (Gentner, 1978, 1982; Macnamara, 1972; Nelson, 1973); therefore, this lexical category was chosen to maintain similarity to natural language acquisition patterns. Additionally, while past research has described the embodied nature of verbs, more recent research has begun to suggest the embodied nature of other lexical categories, including nouns (Saccuman et al., 2006; Arevalo, Perani, Cappa, Butler, Bates, & Dronkers, 2007). In this pilot, 119 German nouns were chosen from the International Picture Naming Project database based on their German word familiarity, frequency, number of syllables and characters. Additionally, reaction times were collected while participants visualized the picture of each noun, visualized interacting with each noun, and actually pantomimed the action associated with each noun in order to establish the average times these different tasks took to complete. From these reaction times, the two experimental training tasks (imagined interaction and picture imagery training) were created to allow for sufficient visualization time in Experiment 2.

**Method**

**Participants.** Participants were 23 right-handed University students. All were English monolinguals between the ages of 18 and 47 (mean age=22.96, SD=6.12; 19 females) and assessed to be healthy and free of psychiatric disorders or developmental disabilities. Participant demographics can be seen in Table 1.

**Procedure.** All participants signed a consent form approved by the University of Houston's Committee for the Protection of Human Subjects. Additionally, background information assessing age, year in school, language history and general health was collected. Expressive English proficiency was assessed with the picture-naming portion of the Woodcock-Munoz Language Proficiency Battery-Revised (Woodcock, 1995). The pilot task required participants to rate 119 black and white pictures of nouns according to four scales. The first scale asked participants to "Rate your familiarity with the object according to how usual or unusual the object is in your realm of experience (the degree to which you come into contact with or think about the concept on a regular basis)" (5=high familiarity…1=low/no familiarity). The second scale asked, "Could you easily mime the action usually associated with this object?" (5= yes/easily…1=no/very hard). The third scale asked, "If you selected that the action associated with this object could be mimed, which body part(s) would you use?" (5=hand(s), 4=feet, 3=mouth, 2=whole body, 1=other). The last scale asked, "How easily does this picture arouse a sensory experience such as a mental picture, interaction, sound, or other physical experience?" (5=very easily…1=not at all). Lastly, participants completed a computer task in which they saw digital pictures of all 119 nouns and were asked to visualize the picture of the object that was just shown on the screen. Participants pressed a button on a computer keyboard when they had completely visualized the picture. Then the task was repeated, however participants were asked to visualize interacting with or using the object. This visualization was similarly timed with a computerized response. The task was completed a third time, however, participants were asked to actually pantomime an action associated with that object. These actions were additionally timed. The orders of these three tasks were counterbalanced and randomized across subjects so that some participants visualized the picture first, others visualized the interaction first, while still others pantomimed the action first to avoid order effects.

**Data analysis.** Ratings on all four scales were used to select 60 nouns that had the highest familiarity, manipulability, and sensory experience scores. These metrics were chosen to ensure that words used in the imagery training paradigms would be easily visualized. These 60 words were randomized using a computer program and split into two groups of 30 words to be learned using different mental imagery techniques in Experiment 2. Using SPSS, univariate ANOVAs were used to assess differences between the two groups of training words on frequency, number of syllables, object complexity, familiarity, ability to mime the action associated with the word, and sensory experience it evoked. The remaining 59 German nouns were used as novel word controls in the experimental fMRI task but not in training; univariate ANOVAs were used to assess differences between the 60 experimental words and 59 novel control words as well, to determine if there were differences in frequency, number of characters, or number of syllables between the two groups. Images of the novel words were never presented to the participants, so these words' object familiarity, manipulability, and sensory scores did not affect the fMRI task. Since auditory presentations of novel words would only be given in the second experiment's fMRI task, the only characteristics of concern were word frequency, number of characters, and number of syllables to ensure they had similar phonological characteristics as the words presented during training.

## Results

**Group 1 and group 2 training words.** From the independent ratings, 60 nouns were chosen and randomly assigned to two experimental groups. These two groups of words would be taught to participants in Experiment 2. Univariate ANOVAs indicated that there were no significant differences ($p < .05$) between group 1 and group 2 experimental words in terms of number of syllables $F (1,58) = .985$, $p = .325$, number of characters $F (1,58) = .006$, $p = .939$, frequency $F (1,58) = .018$, $p = .895$, object complexity $F (1,58) = .743$, $p = .392$, familiarity $F (1,58) = 1.543$, $p = .219$, manipulability $F (1,58) = 1.498$, $p = .226$, or sensory experience $F (1,58) = .643$, $p = .426$. Therefore, it can be concluded that the German words used in both imagery training paradigms were not significantly different based on these parameters. Word characteristics can be seen in Table 2. A list of all words presented in training can be found in Appendix I.

**Novel control words vs. training words.** The remaining 59 words that were not chosen for training were used as novel German controls in the fMRI task that followed vocabulary training. Since these words were never presented during training, ratings of familiarity, manipulability, and sensory experience were not evaluated. Only the number of syllables, characters, and frequency were compared between the 59 control words and 60 experimental words to ensure they had similar phonological characteristics since auditory presentations of these words were given in the fMRI scanner in Experiment 2. Univariate ANOVAs indicated that there were no significant differences ($p < .05$) between experimental and control words on number of characters $F (1,117) = 2.841$, $p = .095$, or frequency $F (1,117) = .027$, $p = .871$; however, the number of syllables approached significance, $F (1,117) = 3.798$, $p = .054$, but the difference was negligible for the purposes of this study. Table 3 lists the characteristics of experimental training and novel control words.

**Reaction times.** In a computer task, participants were timed while they visualized the picture of an object, visualized interacting with an object, or pantomimed an action they associated with an object. Average reaction times for these tasks can be seen in Table 4. In a univariate repeated-measures ANOVA, Mauchly's test indicated that the assumption of sphericity had been violated, Chi-Square (2) = 9.359, $p = .009$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (Epsilon = .728). The corrected results showed that reaction times were significantly affected by the type of task performed (picture visualization, interactive mental imagery, pantomimed action), F (1.456, 30.574) = 43.907, $p < .001$. Follow-up pairwise comparisons between the three tasks, Bonferroni corrected for multiple comparisons, were all significant (p < .05). Visualizing just a picture was fastest, followed by visualizing an interaction. The task requiring the greatest amount of time was the pantomimed action. The average time for actually performing the action was 2840ms (SD = 181.07). Therefore, when designing the imagery training tasks for Experiment 2, five seconds was allotted to participants to create either static or interactive mental imagery for each German noun. This ensured each participant would have plenty of time to complete the visualization tasks in training, as well as the trials that required pantomimed action to ensure participants were visualizing interacting with objects.

## Experiment 2

### Method

**Participants.** Participants were 30 healthy, right-handed English monolingual adults between the ages of 18 and 45 (mean age=22.60, SD=5.51; 23 female). Five participants were dropped from all analyses due to poor fMRI scan acquisition. All participants were screened for language history and proficiency, educational background, and ability to create mental imagery. Handedness was assessed with the Edinburgh Inventory (Oldfield, 1971). Expressive and receptive English language proficiency were respectively measured using the picture vocabulary and listening comprehension sections of the Woodcock-Munoz Language Proficiency Battery-Revised in English (Woodcock, 1995). Ability to create mental images was assessed using the Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973). All participants were screened to be healthy with no history of psychiatric disorders or developmental disabilities. Participant demographics can be seen in Table 5. Participants gave informed consent on a protocol approved by the University of Houston's Human Subjects Committee and were given the choice of two compensation methods: one hour of course extra credit per hour participated or monetary compensation equal to $100 for completion of the entire study. The monetary compensation procedure motivated participants to complete all four days of training and the fMRI scan by providing more money for the latter days of participation. Participants who only completed the screening portion of the study were compensated with $10 or one hour of extra credit.

**Design.** Using a within-subjects design, each participant completed four vocabulary training sessions split into two parts. The first half of the session required participants to learn a set of 30 nouns by visualizing the exact picture of an object they were just shown in conjunction with a novel German word. The second half of the session required participants to learn a second, different set of 30 nouns by visualizing a first-person interaction with the object shown in a picture presented with a novel

German word. To ensure mental images were formed, participants were questioned by the experimenter throughout the learning period based on the type of imagery they used. During picture imagery training, participants were asked to point to the picture they just visualized on a printed array of images. During interactive mental imagery training, participants were asked to demonstrate the interaction they imagined by pantomiming the action they visualized. If participants could not immediately point to the picture they imagined or pantomime the action quickly, they were reminded to complete the task as instructed and visualize images or interactions to the best of their abilities. Training sessions were randomly counterbalanced so that some participants began with interactive mental imagery training, while others began with the picture imagery training. Additionally, training was randomly counterbalanced across word sets so that some participants underwent interactive imagery training with group 1 words, while others underwent interactive imagery training with group 2 words. After each training session, participants took a paper and pencil matching vocabulary test to assess how many words were learned. After all four training sessions were complete, participants were placed in an fMRI scanner while listening to auditory presentations of all 60 learned words and 59 novel words in order to assess neural correlates of word recognition related to the two training methods. Two weeks later, participants returned to take a follow-up vocabulary test to assess vocabulary retention. Figure 1 shows a graphic representation of the experimental timeline.

**Vocabulary training.**

***Imagined interaction.*** In the imagined interaction training, participants were tasked to learn 30 novel German nouns. During training, the participants were seated in a sound booth in front of a computer screen. They were given instructions on how to perform the interactive mental imagery training. Each trial in the imagined interaction training included the presentation of a picture of an object (noun) for 2.5 seconds, visual and auditory presentation of the German word form for 2.5 seconds, and a blank screen for 5 seconds, during which time participants used interactive mental imagery to map the conceptual representation of that object to a novel word. During this time, the subject was asked to imagine a first-person interaction with the object. Prior to the start of training, participants were given detailed instructions with examples of how to create interactive mental images. At least five test trials were given before the start of training where the participant was asked to pantomime what they were imagining to ensure that they understood the instructions. Throughout the 30-minute training session, participants were randomly asked 10 times by the administrator to pantomime the action they just imagined. Quick response helped ensure that subjects were actually trying to create interactive mental images. All 30 words were presented three times in each training session.

***Picture imagery.*** During the same training session, participants were presented with a second, different set of 30 novel German nouns to learn using a different imagery technique. Before training began, participants were instructed that they were going to use a completely different learning strategy for the next set of words. The general format of the training was the same, but the imagery task was different. Participants were similarly presented with a picture for 2.5 seconds, followed by the written and auditory presentation of the German word form for 2.5 seconds; however, during the 5-second imagery period, participants were asked to simply visualize the exact picture of the object they were just shown. After the instructional period, participants underwent a practice session where they completed at least 5 trials in which they were asked to visualize pictures they were just shown. Throughout the 30-minute training session, participants were randomly asked 10 times by the administrator to point to the picture on a large array of images that they just visualized. Quick response helped ensure that subjects were actually trying to create mental images. All 30 words were presented three times in each training session. Figure 2 illustrates both training protocols.

**Scanning procedure.** On the same day as the last training session, participants were placed into an fMRI scanner to assess the neural correlates of word recognition related to the two training methods. The scans took place at Baylor College of Medicine's Human Neuroimaging Laboratory (HNL). Participants were assessed for claustrophobia, medical conditions, as well as presence of metal in the body. All participants signed Baylor's screening and consent form. Participants were told that they would hear words over a set of headphone during the scan. Their task was to simply watch the computer screen because yes/no questions would appear throughout the scan that they needed to answer using handheld button boxes. The 60 German words learned in both training conditions, 59 novel German words, and periods of dedicated silence were presented in a pseudo-randomized event-related design. Stimuli were presented at a mean of once every 8 seconds. Ten times throughout the scan, a question would appear onscreen that asked participants if they had heard a specific word. This task was included in order to keep participants engaged and awake. The total duration of the functional run was about 30 minutes.

***Acquisition parameters.*** Whole-brain scans were acquired using a 3.0 Tesla head-only Siemens Magnetom Allegra imager. A localizer scan assessed each participant's head position prior to data acquisition with the following parameters: voxel size 2.2X1.1X10 mm, repetition time (TR) = 20ms, TE = 5ms. Next, high-resolution T1-weighted anatomical images were collected using a Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence with voxel size 1.0X1.0X1.0 mm, TR = 1200ms, and TE = 2.93ms reconstructed into 192 slices. Functional images were acquired using a clustered volume acquisition event-related paradigm that silenced the scanner while the auditory stimuli were presented. Both learned words (60) and novel German words (59) were pseudo-randomized and presented at a mean of once every 8 seconds. An interleaved descending gradient recalled, echo-planar imaging (EPI) sequence was used during the functional run with voxel size 3.4X3.4X4.0 mm, TR = 2000ms, TR delay (silent interval) = 1420ms, TE = 40ms, flip angle = 90°, and a 64X64 matrix. During the functional run, 26 axial slices per volume were obtained, with the centermost slice aligned parallel to the anterior commissure and posterior commissure (AC-PC) line. During scanning, both the auditory words and the 10 questions were presented using EPrime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

**Data analysis.**

***Behavioral data.*** During each training session, a matching vocabulary test was given, one after picture imagery training, and one after interactive imagery training. Accuracy on the vocabulary tests was recorded. Split-plot repeated measures analysis of variance was used to assess differences in vocabulary test accuracy because even though only one group of monolinguals was assessed, participants could be split into groups based on order of training (interactive imagery first or picture imagery first) or based on the word group used for each type of training (word group 1 for interactive imagery or word group 2 for interactive imagery). Even though both training order and word groups were counterbalanced, these factors were still included in the model in case either of these factors affected the interaction between training paradigm and vocabulary test session number. Therefore, split-plot repeated measures ANOVA with training order and word group as between subjects factors and training paradigm and test session number as within-subjects factors was used to assess differences in vocabulary test accuracy. Correlations assessed the relationship between native language proficiency and learning outcome. Additional correlations assessed the relationship between imagery scores on the VVIQ and vocabulary test accuracy in order to evaluate the relationship between imaging ability and vocabulary learning.

***Imaging data.*** Imaging data was analyzed in SPM8 (Wellcome Trust Centre for Neuroimaging, London) running on a MATLAB 10.0 (The MathWorks, Inc.) platform. In preprocessing, images were realigned for motion correction, resliced, and corrected for timing. The functional images were coregistered to align the mean functional image with the structural image, segmented, and standardized to a standard MNI (Montreal Neurological Institute) template. Data was spatially smoothed with a 8mm Gaussian Kernel (FWHM).

In first-level analyses, stimulus presentation onsets were calculated for each condition (imagined interaction, picture imagery, novel words) and modeled using the General Linear Model (GLM) against an implicit baseline. Motion estimates from preprocessing were used as covariates of no interest to further control for motion artifact (Johnstone et al., 2006). Word condition contrasts (interactive imagery, picture imagery, novel words) from each participant's first-level analysis were entered into a one-way within-subjects ANOVA at the second level. Pairwise contrasts of interest were created to compare activity from each word learning condition (interactive imagery > picture imagery; picture imagery > interactive imagery; learned > novel; novel > learned).

## Results

**Behavioral analyses.** Thirty participants completed all four training sessions and the fMRI scan; however, due to poor fMRI scan acquisition, five participants were dropped from all analyses. Using G*Power (Faul, Erdfelder, Buchner, and Lang, 2009), an *apriori* power analysis for repeated-measures ANOVA indicated that in order to obtain a power of 0.8 with a medium effect size of 0.25 and an alpha of .05, it was necessary to collect a sample of at least 18 participants to detect an interaction between within- and between-subjects factors in the current design.

***Split-plot repeated-measures.*** In order to assess the first hypothesis (Training Method Hypothesis 1), that both training paradigms would result in learning, a split-plot repeated-measures ANOVA with training order and word group as between subjects factors and training paradigm and test session number as within-subjects factors was used to assess differences in vocabulary test accuracy. Results revealed a main effect of training type on accuracy, F (1, 21) = 5.268, $p = .032$. Additionally, Mauchly's test indicated that the assumption of sphericity had been violated for test session number, Chi-Square (9) = 57.367, $p < .001$; therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (Epsilon = .465). The corrected results showed that accuracy on vocabulary tests was significantly affected by test session number, $F(1.858, 39.025) = 51.500$, $p < .001$. Evaluation of interactions in the model revealed the assumption of sphericity had been violated for the interaction of training type and test session number, Chi-Square $(9) = 54.488$, $p < .001$; therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (Epsilon = .500). The interaction of training type and test session number was only significant when either training order or word group was included in the interaction term. In other words, there was a significant three-way interaction between training type, test session number, and training order, $F(1.999, 41.970) = 3.635$, $p = .035$. Follow-up analyses, Bonferroni corrected for multiple comparisons, revealed a significant interaction between training type and test session number for order 1 (interactive imagery training first), $F(1.841, 25.768) = 5.081$, $p = .016$, such that picture imagery training resulted in higher accuracy only on the first vocabulary test (Figure 3). The interaction was not significant for order 2 in which picture imagery training was given first (Figure 4). Therefore, when picture imagery training is given first, there were no significant differences in accuracy between the two training paradigms over the five test sessions; however, when interactive imagery training was given first, accuracy was lower for words learned through interactive imagery on the first test. These results suggest that learning through picture imagery first somehow primes the participants to do better on interactive imagery training from the beginning, possibly through some kind of scaffolding that results in activation of the picture more during interactive training to facilitate learning.

Lastly, there was a significant three-way interaction between training type, test session number, and word group, $F\ (1.999, 41.970) = 4.138$, $p = .023$. Follow-up analyses, Bonferroni corrected for multiple comparisons, revealed a significant interaction for group 2 words, $F\ (1.915, 21.061) = 6.602$, $p = .006$, such that interactive imagery training was more difficult for group 2 words during the first two training sessions and resulted in lower accuracy scores than picture imagery training (Figure 5). The interaction was not significant for group 1 words (Figure 6). Therefore, these results suggest that only group 2 words were harder to learn through interactive imagery in the first two training sessions. However, the two groups of words did not differ on number of syllables, number of characters, frequency, imageability, complexity, manipulability, or ability to evoke a sensory experience (see Experiment 1 Results Section). As such, some other difference between the word groups may have contributed to the greater difficulty participants had learning group 2 words through interactive imagery.

***Correlations.*** In order to evaluate the third and fourth hypotheses (Individual Differences Hypotheses 3 and 4) that higher English proficiency and better imaging ability would result in higher vocabulary test scores, correlation matrices were evaluated. Partial correlations were calculated between vocabulary test scores and English proficiency (Hypothesis 3), controlling for training order, for both picture imagery and interactive imagery since the order of training was found to affect the accuracy results. A weighted composite proficiency score combining both expressive and receptive language scores was used, as the two measures of the Woodcock-Munoz Battery-Revised were correlated, $r = .561$, $p = .002$. All correlations between proficiency and interactive imagery vocabulary scores were significant at the .05 level (Table 6). Results showed composite proficiency was moderately correlated with tests 1 through 5. Additionally, all correlations between proficiency and picture imagery vocabulary scores were significant at the .05 level. Composite proficiency was moderately correlated with all five vocabulary test scores (Table 7). Therefore, for interactive imagery training and picture imagery training, the higher a learner's first language proficiency, the more vocabulary words were learned and retained. These results suggest that native language skill is related to second language vocabulary learning outcomes.

Partial correlations were also calculated between vocabulary test scores and imaging ability (VVIQ scores), controlling for training order, for both picture imagery and interactive imagery training. No correlations were significant ($p < .05$), suggesting that imaging ability, as assessed by the VVIQ, was not related to the ability to learn new vocabulary in either paradigm.

**Neuroimaging analyses.**

***Whole-brain analyses.*** The three word groups, German words learned through interactive imagery, German words learned through picture imagery, and novel German controls were entered into a within-subjects one-way ANOVA at the second level for whole-brain analyses. All results are at presented at a $p$ threshold of .001, uncorrected, with k > 15 to exclude spurious results (Table 8 and Figure 7).

*Training paradigm activation differences.* Evaluation of brain activity for each of the learning conditions modeled against an implicit baseline gave support for the idea that the neural mechanisms involved in word recognition are directly related to the mode of learning (Hypothesis 2). Upon hearing words in the scanner that were learned through picture imagery, participants activated a broad set of bilateral brain regions associated with visual processing, such as the calcarine gyrus, lingual gyrus, cuneus, and precuneus (Table 8 and Figure 7). This is significant since no visual presentation of the lexical word form or a picture of the word was given during the fMRI scan. Hearing words learned through picture imagery activated the visual representation of the word. Additional precentral motor and postcentral somatosensory areas were also active. Alternatively, auditory recognition of words learned through interactive mental imagery showed considerably less neural activity at the group level. These words activated a much more localized network of brain regions, including the left anterior/middle cingulate cortex (ACC), thalamus, and supplementary motor area (SMA). These areas are more involved in general cognitive control in language and non-language related tasks (Abutalebi and Green, 2007). More specifically, both the ACC and the SMA provide function in cognitive motor control. According to Paus (2001), the ACC has the potential to assist in translating intentions into actions, while the SMA functions in articulation and motor control of language (Hernandez, 2009). Therefore, words learned through picture imagery relied more on basic visual and somatosensory processes, whereas words learned through interactive imagery activated regions associated with cognitive motor control of language, suggesting this learning method may have been more embodied.

Pairwise contrasts were also created to directly compare neural activity related to the two learning paradigms. The interactive imagery > picture imagery contrast showed no significant areas of activation. This is significant considering the second hypothesis predicted motor resonance areas would be active in comparisons of words learned through interactive training with picture training. Considering the more localized and attenuated activity in the imagined interaction > implicit baseline condition, it is not surprising though that no regions survived the cross-comparison of both training paradigms. However, this hypothesis is partially supported, as whole-brain contrasts of words learned through interactive imagery compared to baseline did show activity in predicted motor resonance areas as stated above (ACC, SMA, thalamus).

While the direct comparison of interactive imagery and picture imagery did not yield any significant clusters of activation, the picture imagery > imagined interaction contrast did show significant activity in a bilateral set of regions associated with visual processing. When comparing words learned through the different imagery methods, those trained with picture imagery activated bilateral cuneus, the right lingual gyrus, and the left precuneus (Table 8 and Figure 7). Again, the apparent difference in processing between these two groups of words is related to the activation of lower-level visual processing regions after picture imagery training only.

*Comparisons of novel and learned words.* Evaluation of the individual word conditions against an implicit baseline (learned words > implicit baseline; novel words > implicit baseline) showed subtle differences between learned and unlearned words. Learned words activated a very broad network of regions, including the anterior cingulate cortex, superior frontal and precentral areas, and visual areas. These areas were expected, as previous research has shown the anterior cingulate cortex specifically to be activated after successful word learning (Raboyeau et al., 2004). Novel words on the other hand, elicited very little brain activity. Small areas in the pre and post-central gyrus were active; however, the additional expected activity within the right superior temporal gyrus and right inferior frontal gyrus-- areas recognized in novel word recognition (Wong, Perrachione, & Parrish, 2007)-- was not seen. It is thought that with such a passive auditory task, there was no need for the brain to engage areas utilized during difficult lexico-semantic processing. Had the scanner task required participants to more actively engage and manipulate lexico-semantic information, it is possible that the expected activation of right superior temporal and inferior frontal regions would have been seen.

***Post-hoc imaging analyses.***

*Regressions.* In the whole-brain analyses discussed above, there was a lack of robust activation of motor resonance areas in the direct comparison of words learned through interactive imagery to words learned through picture imagery. While whole-brain comparisons of words learned through interactive imagery versus an implicit baseline did activate the ACC, SMA, and thalamus (Table 8), no regions survived a direct comparison of words learned through interactive imagery and picture imagery. It is possible that individual differences in cognition might have produced greater variability in neural mechanisms recruited to complete the interactive imagery training task compared to the picture imagery training task, thus no group-level differences were seen. With this theory in mind, post-hoc regression analyses were done to examine the brain regions that were recruited based off of individual learning trajectories. In other words, by evaluating brain activity as a function of each person's learning slope (overall improvement over the four training periods) and their retention slope (how much vocabulary information was obtained), we gain a better understanding of the neural mechanisms engaged based on how easily or quickly each participant learned in the different training paradigms. While the sample size was not ideal to maximize power to detect differences in an fMRI regression analysis, significant results were still obtained, and therefore interpreted conservatively. These results merely stand to represent promise for future exploration of individual differences in the prediction of successful foreign word learning.

In the regression of words learned through interactive imagery with learning slope (test scores 1-4), there was a positive correlation between slope and activity in the right caudate, right precuneus, right insula, and left cingulate (Table 9 and Figure 8). Therefore, the higher the learning slope (i.e. the greater overall improvement), the more regions involved in cognitive control (caudate, cingulate), memory (insula), and visuospatial processing (precuneus) were activated. In the regression of words learned through picture imagery with the learning slope (test scores 1-4), there was a positive correlation between slope and activity is visual processing regions, such as bilateral superior occipital gyrus (BA 17) and right middle occipital gyrus. Therefore, participants who started out scoring low on picture imagery vocabulary tests relied more heavily on visual processing regions, while those participants that started out scoring low on interactive imagery tests required a broader set of regions involved additionally in cognitive control.

In the regression of words learned through interactive imagery with retention slope (test scores 4-5), there was a negative correlation between slope and activity in bilateral caudate, bilateral cingulate, left SMA, right precuneus, and left postcentral gyrus. These results suggest that participants who lost a significant amount of vocabulary information two weeks after training relied more heavily on cognitive control and motor processing regions during recognition of words learned through interactive imagery right after training. This is indicative of more effortful processing, which suggests the words were not fully entrenched in the lexicon yet. In the regression of words learned through picture imagery with retention slope (tests 4-5), there was a negative correlation between slope and activity in visual processing regions. There was robust activity in left middle occipital gyrus and right superior occipital gyrus, in addition to activity in left superior frontal gyrus, anterior cingulate, postcentral gyrus, and lingual gyrus. These results suggest that the larger the drop in vocabulary retention, the more participants relied on both visual and cognitive control regions when processing words learned through picture imagery right after training. Again the pattern persists that those who struggled in picture imagery training relied more heavily on basic visual processing areas, whereas those that struggled in interactive imagery relied on a broader set of cognitive and motor control regions. Taken together, these results suggest that there is possibly more variability in neural recruitment across subjects in the interactive imagery training, as this is an unfamiliar and therefore more difficult method of learning vocabulary. However, all analyses of words learned through picture imagery led back to visual processing areas. This is consistent with what is expected of words that were learned through visual paired-associate learning. Participants may more easily retrieve vocabulary information through more direct connections in the lower-level sensory systems, whereas higher-level cognitive motor control circuits are recruited during interactive imagery training, suggesting that strategies for learning may be more variable between subjects. This may explain why no significant activity survived group level comparisons of words learned through interactive imagery compared to words learned through picture imagery.

*Training order.* Despite controlling for differences between word groups and counterbalancing training orders, behavioral analyses still showed a significant three-way interaction between training type, test session number, and training order. Therefore, whole-brain fMRI analyses that accounted for effects of training order were also done to see if this behavioral effect transferred over to the neural response to training as well. Whole-brain analyses were done as a 2 x 3 ANOVA, with training order as the between-subjects factor (interactive imagery training first or second) and word type as the within-subjects factor (words learned through interactive imagery, picture imagery, and novel controls). For order 1, in which interactive imagery training was given first, hearing words learned through interactive imagery resulted in activation of the left superior temporal gyrus, left SMA, left temporoparietal junction, right middle occipital, and right supramarginal gyrus (Table 10 and Figure 9). When the order of training was reversed and picture imagery was utilized first, words learned through interactive imagery failed to show activation at the .001 threshold. When a more liberal threshold was used, small areas of activation in the inferior parietal cortex appeared. These results suggest that when picture imagery training is given first, motor resonance areas are no longer activated for words learned through interactive imagery training.

Similarly, the brain regions recruited to process words learned through picture imagery were also influenced by the order of training. For order 1, in which interactive imagery training was given first, words learned through picture imagery activated a small area of the right postcentral parietal cortex, near the supramarginal gyrus, a multisensory integration area. However, when picture imagery training was given first, words learned through picture imagery activated the right calcarine gyrus and left posterior cingulate/precuneus, areas involved in visual processing, as well as the right middle cingulate cortex. Therefore, when picture training is given first, recognition of words learned through pictures activates visual processing regions, whereas if interactive imagery training is given first, picture imagery word recognition relied more on multisensory integration areas. These results suggest that the type of training done first can influence the type of psychological processing engaged in and the concomitant neural processes that are recruited.

## Discussion

The purpose of this study was to investigate whether mental imagery could create a more embodied experience in order to facilitate adult foreign language vocabulary learning. Functional magnetic resonance imaging was used to map the activation differences between two vocabulary training paradigms that used different types of mental imagery, picture imagery training and interactive imagery training. It was hypothesized that the interactive mental imagery training paradigm would facilitate multisensory integration of semantic information more so than picture imagery training, thus strengthening word learning, which would be seen in the activation of motor resonance areas in the brain during recognition of newly learned words and higher accuracy scores after training.

**Training paradigm effects.** After evaluation of the behavioral data, it can be concluded that interactive imagery training did not actually facilitate vocabulary learning beyond picture imagery training as hypothesized; however, both learning methods did result in significant vocabulary gains. Therefore, the first training method hypothesis was only partially supported. Hypothesis one stated that both training paradigms would result in learning, but that interactive imagery training would facilitate greater gains in vocabulary learning and retention. Split-plot repeated measures analysis of variance revealed that there was a significant three-way interaction between training type, test session, and training order. When interactive imagery training was given first, vocabulary scores were lower for words learned through interactive imagery on the first test; however, when picture imagery training was given first, there were no significant differences in accuracy between training paradigms. This suggests that learning with simpler picture visualization first inflates learning through interactive imagery later, but only temporarily. It is possible that by engaging in picture training first, there is a scaffolding effect, whereby active engagement of the visual representation of a word carries over to interactive training, emphasizing this link and thus facilitating word learning.

Alternatively, it is possible that simply practicing visualization, regardless of whether it was static or interactive imagery, primed the participants to do better in the interactive imagery training simply because they had experience creating mental images. Campos, Gomez-Juncal, and Perez-Fabello (2008) examined the effects of experience in mental imagery formation and found that participants who had experience using mental imagery rated imagery vividness higher than those without experience. In their study, Spanish secondary education students were shown sentences and asked to create mental images describing the situation. Having previous experience creating imagery increased their vividness ratings of each sentence scenario, but so did having previous trials of imagery formation; therefore, later trials showed higher imagery vividness scores than earlier trials, suggesting that practice forming mental images increased ability to perform well on the task. It is possible that a similar effect has occurred in this study; by having 30 minutes of practice forming mental images of static pictures, participants are then able to better form interactive mental images. It is possible then that there is not a scaffolding effect whereby the picture representation is activated more strongly, but that a simple practice effect is present.

Further evaluation of these two possible explanations of the order effects yielded additional support for the scaffolding hypothesis found in the neuroimaging results. Whole-brain analyses that evaluated the effects of training order on brain activation in each of the three word conditions (picture imagery, interactive imagery, novel) were conducted and revealed significant effects of training order. When interactive imagery training was given first, words learned through this method activated the SMA, superior temporal gyrus, middle occipital lobe, and regions of the temporoparietal junction. The SMA is responsible for the control and coordination of movement, and is known to contain mirror neurons responsible for motor resonance (Fischer and Zwaan, 2009; Hernandez, 2009). Additionally, areas surrounding the temporoparietal junction such as the supramarginal and angular gyrus are involved in sensory integration (Jeong et al., 2010). Together, these regions were hypothesized to be involved in embodied cognition and active in recognition of words learned through interactive mental imagery. Activation of these regions shows that participants who engaged first in interactive training were most likely forming interactive mental images, as recognition of these words in the scanner resulted in motor resonance. However, the neural activation pattern changed when picture imagery training was given first. When picture imagery training preceded interactive training, recognition of words learned through interactive imagery no longer resulted in motor resonance (Table 10 and Figure 9). When a more liberal *p* threshold was applied, multisensory association areas such as the supramarginal and angular gyrus were significant; however, activation of motor regions was attenuated. It is possible that recognition of these words activated the object representation more than the associated action for these words due to the scaffolding effect that picture training had. This is evidenced by the sole activation of the inferior parietal region, which was expected to be activated in the picture imagery learning condition as well (Hypothesis 2) due to the integration of verbal, visual, and auditory information (Jeong et al., 2010).

**Individual differences in learning.** From the data described above, it is clear that there were both behavioral and neural differences between training methods, particularly when the order of training was taken into account. However, further evaluation of both behavioral and imaging data revealed that in addition to training, individual differences also contributed to variation in learning outcomes. Correlational analyses showed that weighted composite English proficiency was in fact correlated with both successful picture imagery training and interactive imagery training. This suggests that those individuals that have naturally high native language proficiency are better at foreign vocabulary learning, even as adults. This conclusion falls in line with previous literature that suggests native language skills can be successful predictors of foreign language learning (Sparks, Ganschow, Patton, Artzer, Siebenhar, and Plageman, 1997; Meschyan and Hernandez, 2002; Sparks, 2012).

Additionally, it was originally predicted that better imagers would excel at vocabulary learning, particularly in interactive imagery training. However, correlational analyses evaluating the relationship between imaging ability as assessed by the Vividness of Visual Imagery Questionnaire (VVIQ) and vocabulary performance did not support this hypothesis. There were no significant correlations between VVIQ and accuracy when order of training was controlled for. There are two possible explanations for this result. First, it is possible that there is not a strong relationship between the ability to form vivid mental images and the ability to learn vocabulary. However, it is also possible that because the VVIQ is not an all-encompassing measure of imaging ability, that other aspects of imaging ability may still be correlated with vocabulary learning. In their review of the imaging literature, McAvinue and Robertson (2006) describe the most commonly used subjective and objective measures of imaging ability.  According to the authors, it is generally accepted that subjective measures of imaging ability often evaluate self-reported measures of vividness, control, and preference, whereas objective measures often evaluate spatial ability. Unfortunately, McAvinue and Robertson (2006) and much previous literature (Hiscock, 1978; Di Vesta, Ingersoll, & Sunshine, 1971, Neisser, 1970) suggests that these subjective and objective measures are often unrelated. Therefore, it is possible that while imaging ability as assessed by a self-reported measure of vividness (VVIQ) is not correlated with vocabulary performance, a more objective measure of spatial ability such as those used in Kosslyn's (2001; 2003) studies might be.

Moreover, as the individual differences hypotheses were partially supported in the behavioral data, post-hoc neuroimaging analyses were done to preliminarily begin to examine individual variability in neural responses to the different learning methods. FMRI regression analyses revealed that in interactive imagery learning, the higher the learning slope (lower starting score and greater improvement), the more regions involved in visual processing and cognitive motor control were relied upon during word recognition. Specifically, the right precuneus, right caudate, right insula, and left cingulate were active in those participants who struggled to learn at first but excelled later in training. Similarly, cognitive control regions were active in those participants who failed to retain as much vocabulary information after two weeks. However, in picture imagery learning, higher learning slopes as well as lower retention slopes were correlated with activity in lower-level visual processing regions in the occipital gyrus. These results suggest that a broader set of cognitive and motor control regions were recruited in interactive learning, whereas more localized activity of visual processing regions supported picture learning in those participants that first struggled to learn.

In almost all of the neuroimaging analyses of picture imagery training, there seems to be more homogeneous activity of basic visual processing areas (Figure 7, Figure 8, and Figure 9); This is consistent with what is expected of words learned through a picture-associated learning paradigm; however, in the analyses of interactive imagery training, there seems to be more distributed activity in motor, sensory integration, and cognitive control regions. This greater variability in neural recruitment during interactive imagery training helps to explain why no regions survived group-level comparisons of words learned through interactive imagery to picture imagery; the individual variability washed out group-level effects. It is therefore clear from both the behavioral correlations and the imaging data that individual variability, possibly more than training method, has the potential to influence learning outcomes.

Along with the behavioral analyses that showed interactive imagery learners struggled at the start of training, the individual variability in neural recruitment during interactive imagery training suggests that this mode of learning was more difficult for adult learners, and therefore required participants to adopt varying strategies or neural mechanisms to cope. The particular design of this study was not optimized to further examine the potential for individual differences in processing. It was predicted that there would be more robust differences in learning based on the mode of learning, with individual differences explaining any other variability in the results. This was not the case. There were very few differences in performance based on the training method used, whereas the influences of individual differences were great at both the behavioral and neural levels. A much larger sample size, as well as the addition of additional predictors of vocabulary learning success need to be included in future studies, as it appears that these differences may drive learning outcomes more so than the method of instruction. While vocabulary performance was affected by the order of the training paradigms, by session two, these differences were attenuated in later training sessions. Therefore, future studies investigating ways to achieve successful foreign word learning may benefit from focusing on individual differences in learning as opposed to investigating how to make vocabulary learning more embodied.

**Conclusions**

In this study, we investigated whether mental imagery could be used to create an embodied experience that might facilitate second language vocabulary learning. We evaluated vocabulary learning by measuring performance on two types of imagery-based training tasks, as well as mapped the neural correlates of these learning methods. Additionally, individual difference predictors of vocabulary performance were also evaluated-- native language proficiency and imaging ability. We found that while the interactive imagery paradigm did produce more embodied responses to words learning through this method (activation of the SMA, ACC, and thalamus), it did not result in significant gains in vocabulary learning above the picture imagery method. Additionally, it seems that when picture imagery training was conducted before interactive training, there was a scaffolding effect during the first training session, whereby learning to activate the representation of an object aided the facilitation of interactive training. Therefore, it seems that picture imagery training actually facilitated vocabulary learning above and beyond interactive training—the opposite of what was hypothesized. Lastly, we found that while imaging ability did not predict vocabulary learning success, native language proficiency did. Analysis of the neural mechanisms driving individual learning differences gave further support that individual variability and not just instructional method may drive successful foreign vocabulary acquisition.

Overall, this study contributes to a growing body of literature that assesses factors that might aid second language vocabulary learning. While there were several limitations in the study design that prevented the further investigation of individual predictors of successful vocabulary learning (small sample size, complex within-subjects design, lack of cognitive pretests), we were able to discover the foundational information that interactive mental imagery, while utilizing more embodied neural mechanisms, did not contribute to more successful language learning than picture imagery. In fact, picture imagery training actually primed learners to perform better during interactive imagery training. Additionally, individual differences, both behavioral and neural were able to characterize variability in learning outcomes. Taken together, these results will help us design future studies that examine other predictors of successful second language acquisition, which could potentially influence the way second language instruction is conceptualized.

## References

Abutalebi, J., and Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242-275.

Al-Seghaer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232.

Arevalo, A., Perani, D., Cappa, S.F., Butler, A., Bates, E., & Dronkers, N. (2007). Action and object processing in aphasia: from nouns and verbs to the effect of manipulability. *Brain and Language*, *100*, 79–94.

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

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

Bergen, B., Lau, T. C., Narayan, S. Stojanovic, D., & Wheeler, K. (2010). Body part representations in verbal semantics. *Memory & Cognition, 38*(7), 969-981.

Campos, A., Gomez-Juncal, R., & Perez-Fabello, M. J. (2007). Experience in imagery and imagery vividness. *Imagination, Cognition, and Personality, 27*(4), 337-348.

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.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992).

Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Di Vesta, F. J., Ingersoll, G., & Sunshine, P. (1971). A factor analysis of imagery tests. *Journal of Verbal Learning and Verbal Behavior, 10,* 471-479.

Di Vesta, F. J., & Ross, S. M. (1971). Imagery ability, abstractness, and word order as variables in recall of adjectives and nouns. *Journal of verbal learning and verbal behavior, 10,* 686-693.

Dufva, M., & Voeten, M. J. M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholoinguistics, 20*, 329-348.

Ernest, C. H., & Paivio, A. (1969). Imagery ability in paired-associate and incidental learning. *Psychonomic Science, 15,* 181-182.

Ernest, C. H., & Paivio, A. (1971). Imagery and verbal associative latencies as a function of imagery ability. *Canadian Journal of Psychology, 25*(1), 83-90.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2009). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied languages: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*(6), 825-850.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development, 49,* 988-998.

Gentner, D. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj II (ed) Language development, volume 2: Language, thought and culture. Hillsdale, N.J.: Lawrence Erlbaum, 1982. P. 301-334.

Glenberg, A. M., Goldberg, A. B., & Zhu, X. (2011). Improving early reading comprehension using embodied CAI. *Instructional Science, 39,* 27-39.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*, 424–436.

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

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*, 1-26.

Hernandez, A. E. (2009). Language switching in the bilingual brain: What's next? *Brain & Language, 109*(2-3)*,* 133-140.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences, 9*(5), 220-225.

Hiscock, M. (1978). Imagery assessment through self-report: What do imagery questionnaires measure? *Journal of Consulting and Clinical Psychology, 46*(2), 223-230.

James, K. H. & Maouene, J. (2009). Auditory verb perception recruits motor systems in the developing brain: An fMRI investigation. *Developmental Science,* 26-34.

Jeong, H., Sugiura, M., Sassa, Y., Wakusawa, K., Horie, K., Sato, S., et al. (2010). Learning second language vocabulary: Neural dissociation of situation-based learning and text-based learning. *NeuroImage, 50*, 802-809.

Johnstone, T., Ores Walsh, K. S., Greischar, L. L., Alexander, A. L., Fox, A. S., Davisdon, R. J., and Oakes, T. R. (2006). Motion correction and the use of motion covariates in multiple-subject fMRI. *Human Brain Mapping, 27*, 779-788.

Jones, L. & Stuth, G. (1997). The uses of mental imagery in athletics: An overview. *Applied & Preventive Psychology, 6*(2), 101-115.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. Journal of Memory & Language, 43, 508-529.

Keysers, C., Kohler, E., Umilta, M. A., Fogassi, L., Nanetti, L., & Gallese, V. (2003). Audio-visual mirror neurons and action recognition. Experimental Brain Research, 153, 628–636.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*, 635-642.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin, 129*(5), 723-746.

Lutz, K. A, & Lutz, R. J. (1977). Effects of interactive imagery on learning: Application to advertising. *Journal of Applied Psychology, 62*(4), 493-498.

Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review, 79,* 1-13.

Marks, D.F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*, 17-24.

McAvinue, L. P., & Robertson, I. H. (2006). Measuring visual imagery ability: A review. *Imagination, Cognition and Personality, 26*(3), 191-211.

Meschyan, G., & Hernandez, A. (2002). Is native-language decoding skill related to second language learning? *Journal of Educational Psychology, 94*(1), 14-22.

Neisser, U. (1970). Visual imagery as process and as experience. In J.S. Antrobus (Ed.), *Cognition and affect.* Boston: Little, Brown.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38,* 149.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron*, *36*, 1211–1220.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, and Winston. (Reprinted 1979; Hillsdale, NJ: Lawrence Erlbaum Associates.)

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

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255–287.

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience, 2*, 417-424.

Pressley, M. (1977). Imagery and children's learning: Putting the picture in developmental perspective. *Review of Educational Research, 47*(4), 585-622.

Pulvermüller, F., Harle, M., & Hummel, F. (2001). Walking or talking: behavioral and neurophysiological correlates of action verb processing. *Brain and Language*, *78*, 143–168.

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

Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Demonet, J. F., Cardebat, D. (2004). Lexical learning of the English language: A PET study in healthy French subjects. *NeuroImage, 22,* 1808-1818.

Saccuman, M. C., Cappa, S. F., Bates, E.A., Arevalo, A., Della Rosa, P., Danna, M., & Perani, D. (2006). The impact of semantic reference on word class: an fMRI study of action and object naming. *NeuroImage*, *32*, 1865–1878.

Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly:Overcoming Learning Difficulties, 2*1(3), 221-238.

Sanz, C. (Ed.). (2005). *Mind and context in adult second language acquisition: methods, theory, and practice*. Washington, D.C.: Georgetown University Press.

Schwartz, D. L. & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology, 25*(1), 116-136.

Semin, G. R., & Smith, R. E. (2008). Embodied grounding: Social cognitive, affective, and neuroscientific approaches. New York, NY: Cambridge University Press.

Shen, H. (2010). Imagery and verbal coding approaches in Chinese vocabulary instruction. *Language Teaching Research, 14*(4), 485-499.

Sparks, R. L. (2012). Individual differences in L2 learning and long-term L1-L2 relationships. *Language Learning, 62*(Suppl. 2), 5-27.

Sparks, R. & Ganschow, L. (2001). Aptitude for learning a foreign language. Annual Review of Applied Linguistics, 21, 90-111.

Sparks, R. L., Ganschow, L., Patton, J., Artzer, M., Siebenhar, D., & Plageman, M. (1997). Prediction of foreign language proficiency. *Journal of Educational Psychology, 89*(3), 549-561.

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency. *Applied Psychollinguistics, 30*, 725-755.

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2011). Subcomponents of second-language aptitude and second-language proficiency. *The Modern Language Journal, 95*(ii), 253-273.

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Do L1 reading achievement and L1 print exposure contribute to the prediction of L2 proficiency? *Language Learning, 62*(2), 473-505.

Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia, 56*(1), 129-160.

Spolsky, B. (1995). Measured Words. Oxford University Press, Oxford.

Symonds, P. (1930). *Foreign Language Prognosis Test*. New York: Teachers College, Columbia University.

Whitfield-Gabrielli, S. (2010). Artifact Detection and QA manual, from http://web.mit.edu/swg/art/art.pdf

Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping, 28*, 995-1006.

Woodcock, R. W. (1995). *Woodcock Language Proficiency Battery - Revised (WLPB-R)*: Riverside Publishing.

Woody, R. H. (2006). Musicians' cognitive processing of imagery-based instructions for expressive performance. *Journal of Research in Music Education, 54*(2), 125-137.

## Table 1. Experiment 1 Demographics

| | Demographics | | | |
|---|---|---|---|---|
| | **n** | **Sex** | **Age** | **Woodcock-Munoz English Vocabulary** |
| **English Monolinguals** | 23 | 4 male/19 female | 22.96 (SD=6.12) | 45.35 (SD=3.93) |

## Table 2. Group 1 and Group 2 Training Word Characteristics

| | Word Characteristics | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **n** | **# Syllables** | **# Characters** | **Frequency** | **Object Complexity** | **Familiarity** | **Manipulability** | **Sensory** |
| **Group 1** | 30 | 1.87 (SD=.346) | 6.17 (SD=1.76) | 2.51 (SD=1.20) | 13,560.27 (SD=4288.56) | 4.73 (SD=.207) | 4.65 (SD=.232) | 4.56 (SD=.167) |
| **Group 2** | 30 | 1.77 (SD=.430) | 6.20 (SD=1.58) | 2.47 (SD =1.17) | 15,056.33 (SD=8482.18) | 4.64 (SD=.315) | 4.57 (SD=.244) | 4.59 (SD=.155) |

## Table 3. Training and Novel Word Characteristics

| | Word Characteristics | | | |
|---|---|---|---|---|
| | **n** | **# Syllables** | **# Characters** | **Frequency** |
| **Experimental Training Words** | 60 | 1.82 (SD=.390) | 6.18 (SD=1.66) | 2.49 (SD=1.17) |
| **Novel Control Words** | 59 | 1.66 (SD=.477) | 5.66 (SD=1.72) | 2.53 (SD=1.33) |

## Table 4. Experiment 1 Reaction Times

| Reaction Times (RT) | | |
|---|---|---|
| | **n** | **RT** |
| **Picture Visualization** | 22 | 1505.23 (SD=1006.57) |
| **Interactive Imagery** | 22 | 1777.10 (SD=868.49) |
| **Pantomimed Action** | 22 | 2840.86 (SD=849.28) |

## Table 5. Experiment 2 Demographics

| Demographics | | | | | | | |
|---|---|---|---|---|---|---|---|
| | **n** | **Sex** | **Age** | **Woodcock-Munoz Vocabulary** | **Woodcock-Munoz Comprehension** | **Weighted Composite Proficiency (out of 100)** | **VVIQ Total Score** |
| **Monolinguals** | 25 | 23 female/2male | 22.60 (SD=5.51) | 47.60 (SD=3.52) | 27.92 (SD=2.96) | 82.64 (SD=6.63) | 27.56 (SD=8.50) |

**Table 6. Correlation Matrix 1**
This table displays the correlation of interactive imagery vocabulary scores and proficiency, controlling for order or training.

| Correlation Matrix | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Proficiency | Int Score 1 | Int Score 2 | Int Score 3 | Int Score 4 | Int Score 5 |
| Proficiency | Correlation | 1 | .742 | .627 | .359 | .702 | .532 |
| | Significance | | .000 | .001 | .042 | .000 | .004 |
| | df | 0 | 22 | 22 | 22 | 22 | 22 |
| Int Score 1 | Correlation | .742 | 1 | .710 | .142 | .781 | .402 |
| | Significance | .000 | | .000 | .254 | .000 | .026 |
| | df | 22 | 0 | 22 | 22 | 22 | 22 |
| Int Score 2 | Correlation | .627 | .710 | 1 | .545 | .721 | .469 |
| | Significance | .001 | .000 | | .003 | .000 | .010 |
| | df | 22 | 22 | 0 | 22 | 22 | 22 |
| Int Score 3 | Correlation | .359 | .142 | .545 | 1 | .362 | .457 |
| | Significance | .042 | .254 | .003 | | .041 | .012 |
| | df | 22 | 22 | 22 | 0 | 22 | 22 |
| Int Score 4 | Correlation | .702 | .781 | .721 | .362 | 1 | .378 |
| | Significance | .000 | .000 | .000 | .041 | | .034 |
| | df | 22 | 22 | 22 | 22 | 0 | 22 |
| Int Score 5 | Correlation | .532 | .402 | .469 | .457 | .378 | 1 |
| | Significance | .004 | .026 | .010 | .012 | .034 | |
| | df | 22 | 22 | 22 | 22 | 22 | 0 |

### Table 7. Correlation Matrix 2

This table displays the correlation of picture imagery vocabulary scores and proficiency, controlling for order of training.

| Correlation Matrix | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Proficiency | Picture Score 1 | Picture Score 2 | Picture Score 3 | Picture Score 4 | Picture Score 5 |
| Proficiency | Correlation | 1 | .446 | .493 | .479 | .476 | .511 |
| | Significance | | .015 | .007 | .009 | .009 | .005 |
| | df | 0 | 22 | 22 | 22 | 22 | 22 |
| Picture Score 1 | Correlation | .446 | 1 | .809 | .576 | .573 | .730 |
| | Significance | .015 | | .000 | .002 | .002 | .000 |
| | df | 22 | 0 | 22 | 22 | 22 | 22 |
| Picture Score 2 | Correlation | .493 | .809 | 1 | .867 | .797 | .731 |
| | Significance | .007 | .000 | | .000 | .000 | .000 |
| | df | 22 | 22 | 0 | 22 | 22 | 22 |
| Picture Score 3 | Correlation | .479 | .576 | .867 | 1 | .857 | .777 |
| | Significance | .009 | .002 | .000 | | .000 | .000 |
| | df | 22 | 22 | 22 | 0 | 22 | 22 |
| Picture Score 4 | Correlation | .476 | .573 | .797 | .857 | 1 | .713 |
| | Significance | .009 | .002 | .000 | .000 | | .000 |
| | df | 22 | 22 | 22 | 22 | 0 | 22 |
| Picture Score 5 | Correlation | .511 | .730 | .731 | .777 | .713 | 1 |
| | Significance | .005 | .000 | .000 | .000 | .000 | |
| | df | 22 | 22 | 22 | 22 | 22 | 0 |

**Table 8. Whole-Brain fMRI Results**

Whole-brain analyses were conducted at a p threshold of .001, uncorrected with a minimum of 15 voxels per cluster. Coordinates are reported in MNI space.

| Condition | p | k | Region | Side | x | y | z |
|---|---|---|---|---|---|---|---|
| Interactive Imagery Words > Implicit Baseline | 0.001 | 32 | Superior frontal gyrus BA6 | R | -18 | 2 | 62 |
| | | 26 | Middle cingulate BA24 | L | -12 | 8 | 40 |
| | | 21 | Thalamus | Bilateral | 0 | -21 | 0 |
| Picture Imagery Words > Implicit Baseline | 0.001 | 303 | Superior frontal BA8 | R | 20 | 22 | 36 |
| | | 410 | Middle frontal BA6 (SMA) | L | -22 | 8 | 40 |
| | | 99 | Precentral BA4 | L | -32 | -16 | 48 |
| | | 39 | Cuneus BA19 | L | -32 | -86 | 30 |
| | | 54 | Postcentral BA5 | L | -30 | -38 | 50 |
| | | 115 | Cuneus | R | 18 | -78 | 32 |
| | | 51 | Postcentral BA2 | R | 36 | -30 | 24 |
| | | 56 | Postcentral BA3 | R | 36 | -26 | 42 |
| | | 73 | Lingual gyrus | R | 12 | -46 | 0 |
| | | 134 | Precuneus | L | -20 | -66 | 28 |
| | | 16 | Calcarine gyrus BA17 | R | 12 | -92 | 12 |
| | | 77 | Lingual gyrus | L | -8 | -54 | 0 |
| All Learned Words > Implicit Baseline | 0.001 | 241 | Precentral BA4 | L | -32 | -16 | 46 |
| | | 258 | Cingulate | L | -16 | 6 | 40 |
| | | 249 | Superior frontal BA8 | R | 20 | 22 | 36 |
| | | 34 | ACC | R | 18 | 44 | 12 |
| | | 83 | Lingual gyrus | R | 12 | -50 | 2 |
| | | 46 | Middle cingulate | R | 14 | -24 | 44 |
| | | 53 | Postcentral BA40 | R | 40 | -26 | 44 |
| | | 71 | Middle frontal BA6 | R | 26 | -6 | 40 |
| | | 36 | ACC | L | -16 | 26 | 14 |
| | | 28 | Cuneus | R | 16 | -86 | 32 |
| | | 21 | Precuneus | L | -2 | -48 | 46 |
| | | 19 | Precuneus | R | 28 | -46 | 44 |
| Novel Words > Implicit Baseline | 0.001 | 18 | Postcentral | R | 12 | -34 | 70 |
| | | 87 | Precentral BA4 | R | 54 | -6 | 30 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | 35 | Postcentral BA2 | L | -44 | -22 | 50 |
| Learned Words > Novel Words | 0.001 | 14 | Medial frontal BA9 | R | 20 | 44 | 10 |
| Picture Imagery > Imagined Interaction | 0.001 | 182 | Posterior cingulate/Precuneus | L | -8 | -50 | 12 |
| | | 38 | Cuneus BA18 | L | -22 | -78 | 20 |
| | | 15 | Cuneus | R | 18 | -78 | 32 |
| | | 20 | Lingual gyrus | R | 12 | -42 | -2 |
| Picture Imagery Words > Novel Words | 0.001 | 20 | Middle frontal BA32 | L | -16 | 20 | 36 |
| | | 25 | Middle frontal BA8 | R | 24 | 24 | 34 |

**Table 9. fMRI Regression Results**

Regression analyses are reported at a p threshold of .001, uncorrected with a minimum of 15 voxels per cluster. Coordinates are reported in MNI space.

| Regressions | p | k | Area | Side | x | y | z |
|---|---|---|---|---|---|---|---|
| Interaction & Learning Slope (test 1-4) | 0.001 | 166 | Precuneus BA 23 | R | 2 | -56 | 14 |
| Positive Correlation | | 64 | Caudate | R | 18 | -10 | 16 |
| | | 52 | Precuneus BA 31 | R | 14 | -44 | 22 |
| | | 48 | Cingulate | L | -4 | -6 | 34 |
| | | 44 | Caudate | R | 20 | 16 | 8 |
| | | 23 | Superior temporal gyrus/Insula | R | 46 | -18 | -8 |
| Interaction & Retention Slope (test 4-5) | 0.001 | 140 | Caudate | R | 18 | 20 | 10 |
| Negative Correlation | | 100 | Caudate | L | -16 | 16 | 10 |
| | | 59 | Postcentral gyrus | L | -48 | -34 | 56 |
| | | 50 | Precuneus | R | 18 | -60 | 42 |
| | | 49 | SMA | L | -6 | 22 | 54 |
| | | 44 | Cingulate | R | 12 | 12 | 22 |
| | | 32 | Cingulate | R | 12 | 30 | 36 |
| | | 31 | Anterior cingulate | L | -12 | 26 | 18 |
| | | 24 | White matter | R | 20 | -30 | 16 |
| | | 17 | Cingulate BA 32 | R | 20 | 12 | 38 |
| Picture & Learning Slope (test 1-4) | 0.001 | 29 | Superior occipital gyrus area 17 | L | -14 | -88 | 10 |
| Positive Correlation | | 21 | Middle occipital gyrus | L | -24 | -84 | -2 |
| | | 20 | Superior occipital gyrus | R | 22 | -82 | 18 |
| Picture & Retention Slope (test 4-5) | 0.001 | 353 | Left middle occipital gyrus | L | -18 | -90 | 10 |
| Negative Correlation | | 305 | Superior occipital gyrus | R | 24 | -70 | 20 |
| | | 91 | Superior frontal BA10 | L | -20 | 48 | 14 |

| | | | | | |
|---|---|---|---|---|---|
| 88 | ACC | L | -2 | 18 | 14 |
| 74 | ACC | L | -4 | 38 | 26 |
| 69 | Lingual gyrus | R | 14 | -72 | -6 |
| 59 | Postcentral gyrus BA3 | L | -30 | -18 | 38 |
| 15 | Cingulate BA32 | R | 18 | 26 | 20 |

**Table 10. fMRI Order Effects**

Results are reported at a p threshold of .001, uncorrected with a minimum of 10 voxels per cluster. Coordinates are reported in MNI space.
Order 1 (O1) = Interactive Imagery Training First
Order 2 (O2) = Picture Imagery Training First

| Conditions | p | Area | Side | k | x | y | z |
|---|---|---|---|---|---|---|---|
| O1 > O2 (Picture Imagery Words) | 0.001 | postcentral parietal | R | 14 | 64 | -12 | 14 |
| O1 > O2 (Interactive Imagery Words) | 0.001 | superior temporal gyrus | L | 150 | -46 | -15 | -2 |
| | | SMA | L | 53 | -3 | 4 | 56 |
| | | supramarginal gyrus BA 40 | R | 26 | 62 | -13 | 15 |
| | | middle/superior occipital BA 19 | R | 24 | 34 | -87 | 20 |
| | | superior temporal gyrus BA 22/temporoparietal junction | L | 22 | -63 | -41 | 9 |
| O1 > O2 (Novel Words) | 0.001 | rolandic operculum | R | 331 | 58 | -9 | 16 |
| | | precentral BA4/inferior frontal | L | 142 | -54 | -2 | 19 |
| | | Inferior frontal BA9 | L | 109 | -50 | 22 | 28 |
| | | inferior parietal BA40 | L | 84 | -40 | -33 | 24 |
| | | superior temporal BA22/temporoparietal junction | L | 22 | -60 | -40 | 4 |
| O1 > O2 (Learned Words) | 0.001 | Rolandic Operculum | R | 73 | 63 | -13 | 12 |
| | | superior temporal | L | 12 | -62 | -27 | 6 |
| O2 > O1 (Picture Imagery Words) | 0.001 | posterior cingulate/precuneus BA31 | L | 16 | -16 | -55 | 16 |
| | | middle cingulate | R | 10 | 14 | -28 | 45 |
| | | calcarine gyrus | R | 10 | 15 | -48 | 11 |
| O2 > O1 (Interactive Imagery Words) | 0.001 | None | | | | | |
| | 0.005 | Supramarginal/angular gyrus | L | 46 | -49 | -53 | 34 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Supramarginal/angular gyrus | R | 30 | 54 | -53 | 29 |
| O2 > O1 (Novel) | 0.001 | cerebellum anterior lobe, culmen | R | 11 | 19 | -33 | -18 |
| O2 > O1 (Learned) | 0.001 | None | | | | | |

**Figure 1. Experiment Timeline**
There were four days of vocabulary training. Each session was split into two parts for the two types of imagery training. There was a vocabulary test after each training session, as well as two weeks after training was completed. There was one fMRI scan after the fourth day of training was complete.

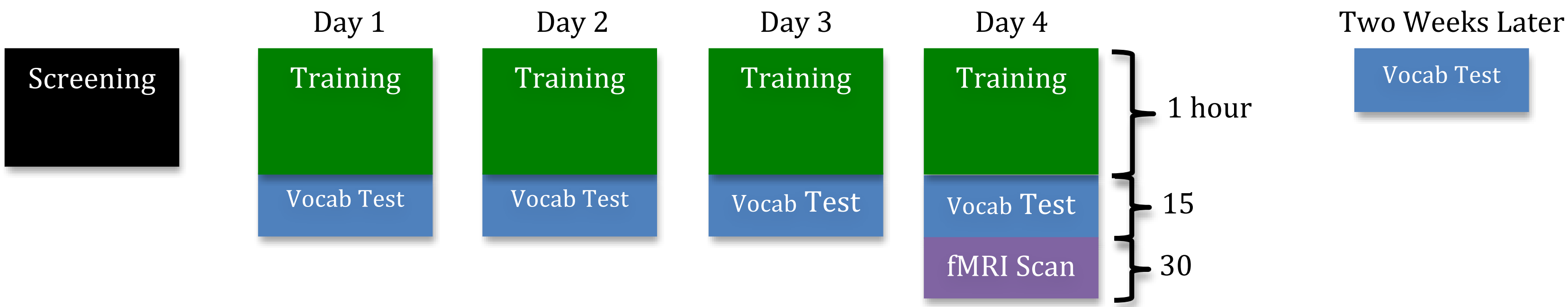


**Figure 2. Training Paradigms**

This figure shows one trial (one word) from the training sessions. The presentation of the stimuli is the same between both training conditions. The only difference is the set of words presented and the imagery task during the time the blank screen is presented. The 30 words used in the imagined interaction training are different than the 30 words used in the picture imagery training.

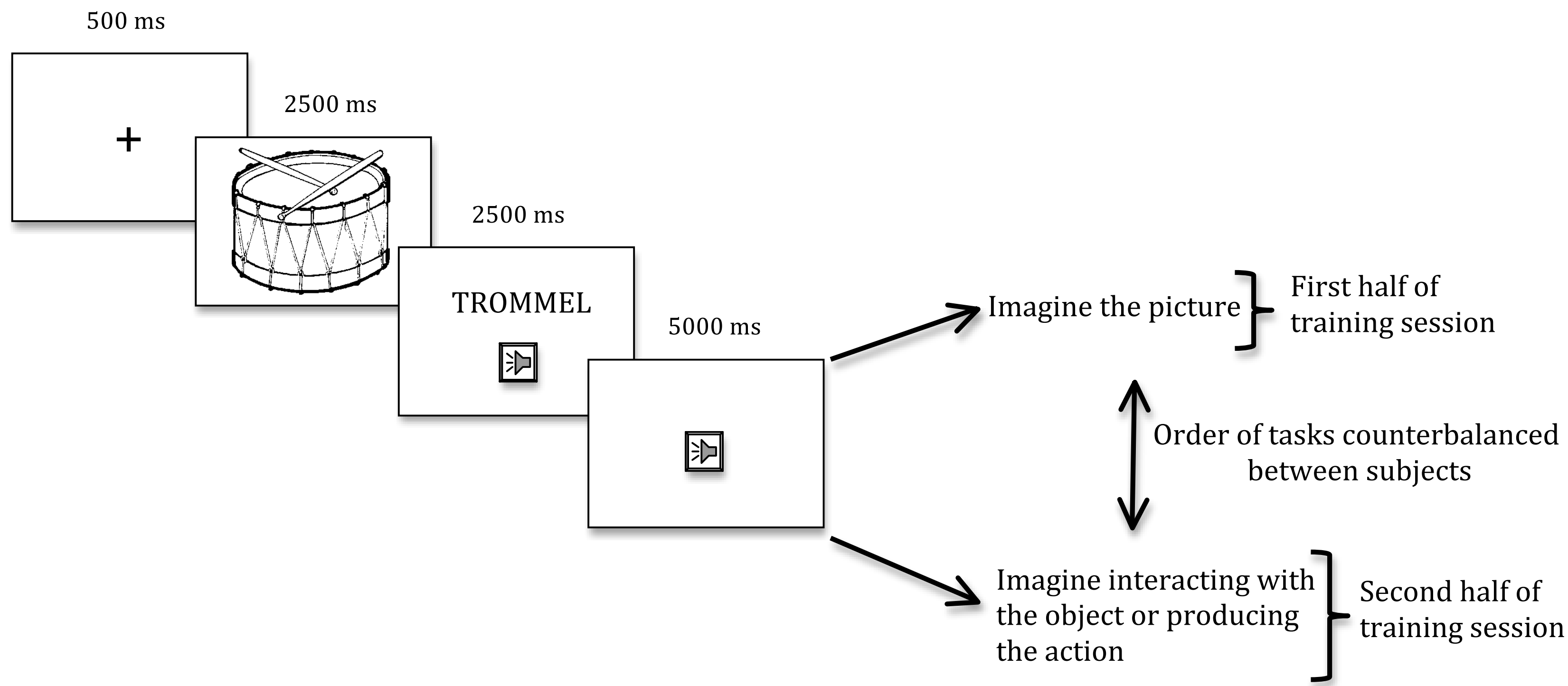


**Figure 3. Order 1 Interaction of Training Type and Test Session Number**
In order 1, interactive imagery training was given first, followed by picture training. Significant differences are marked with as asterisk. The green line represents interactive imagery training, whereas the blue line represents picture imagery training.

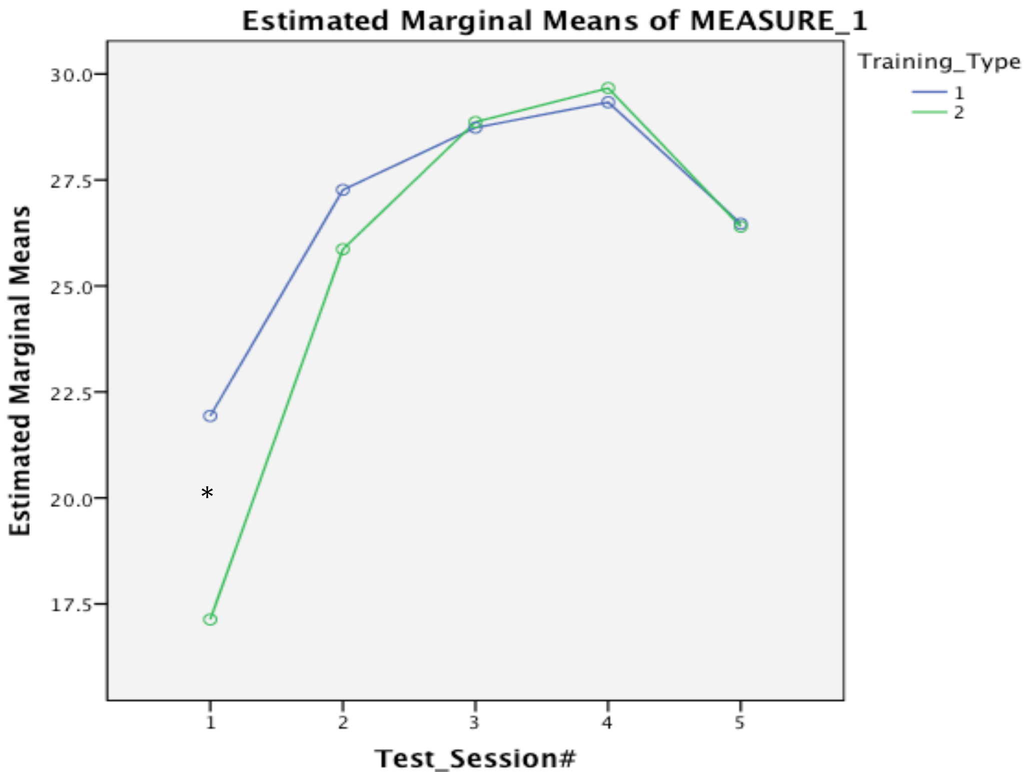


**Figure 4. Order 2 Interaction of Training Type and Test Session Number**
In order 2, picture imagery training was given first, followed by interactive imagery training. There were no significant differences between training methods when picture imagery training was given first. The green line represents interactive imagery training, whereas the blue line represents picture imagery training.

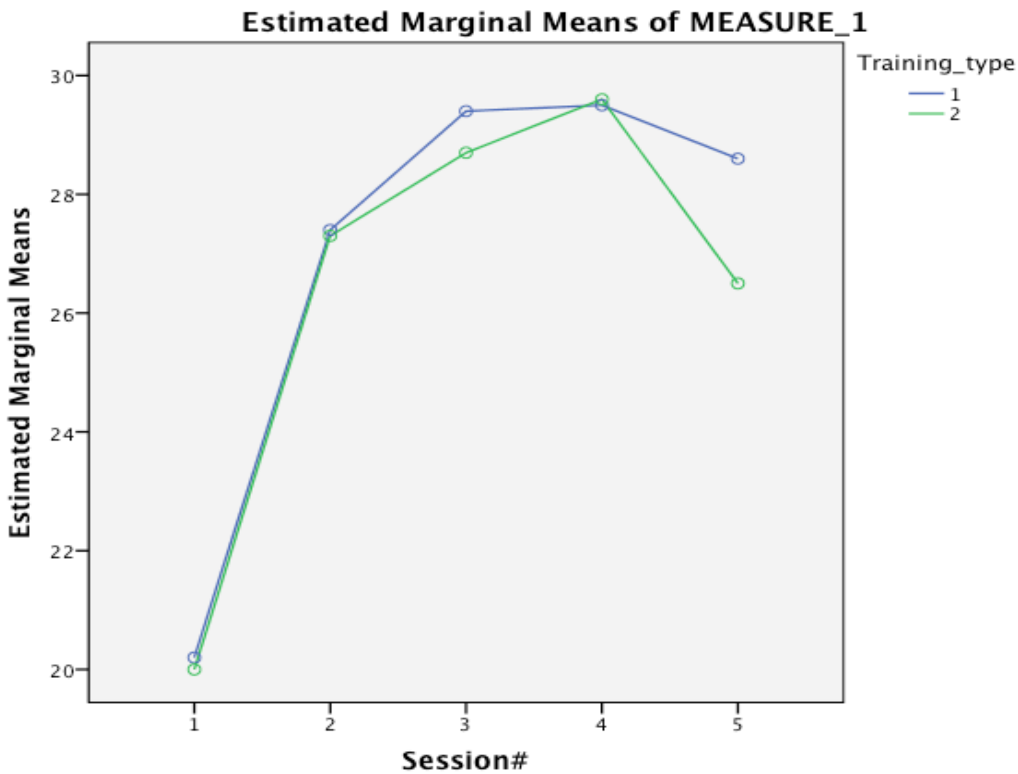


**Figure 5. Word Group 2 Interaction of Training Type and Test Session Number**
Significant differences are marked with an asterisk. Interactive imagery training was more difficult for group 2 words for the first two training sessions and resulted in lower accuracy scores than picture imagery training. The green line represents interactive imagery training, whereas the blue line represents picture imagery training.

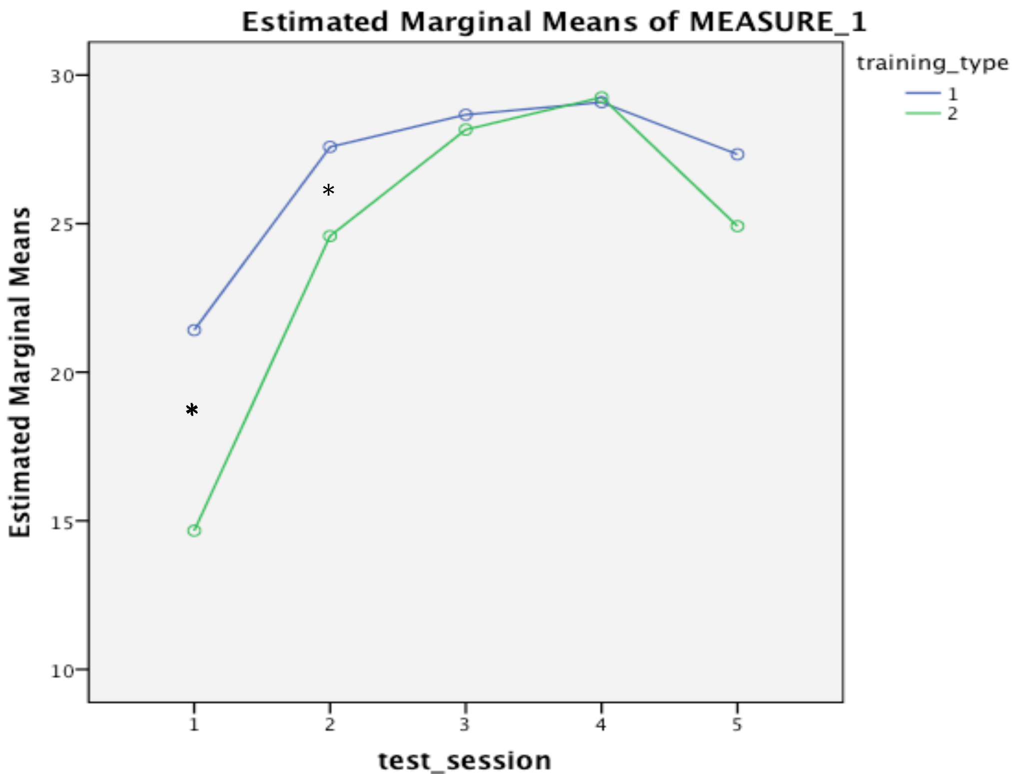


**Figure 6. Word Group 1 Interaction of Training Type and Test Session Number**
There were no significant differences between training methods for group 1 words. The green line represents interactive imagery training, whereas the blue line represents picture imagery training.

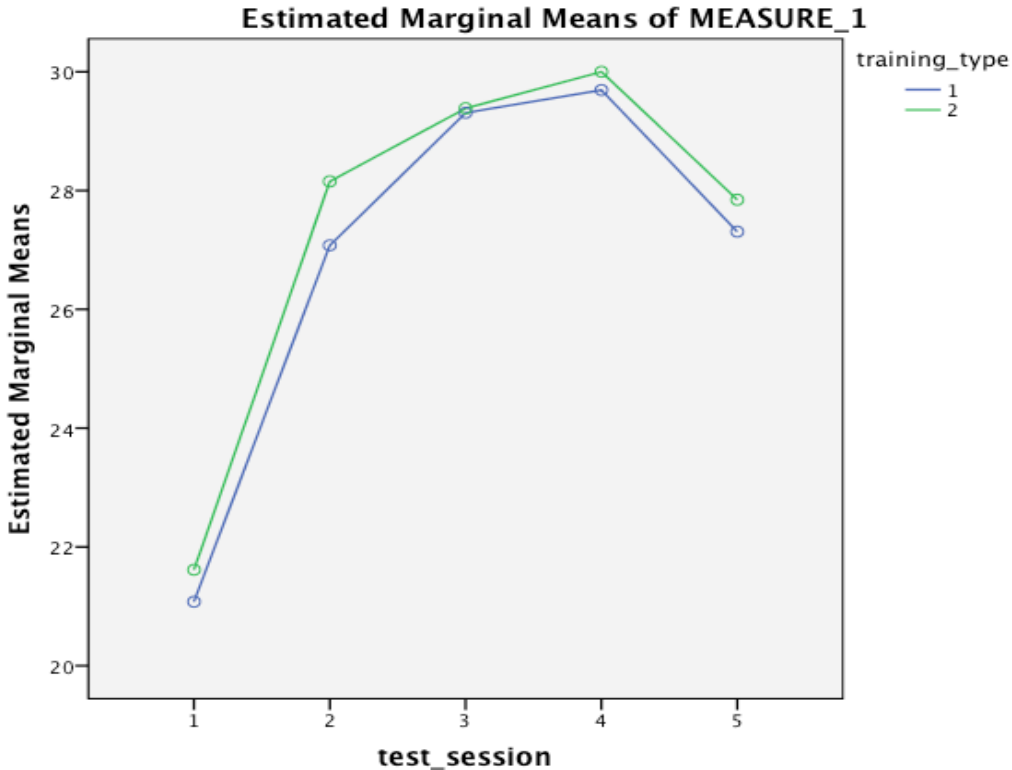
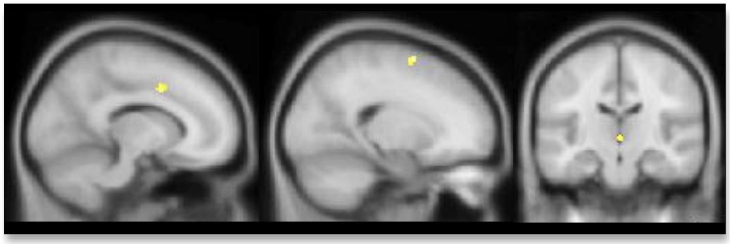
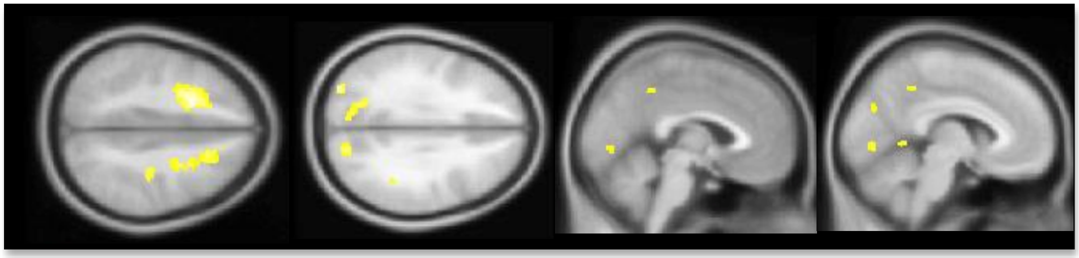


**Figure 7. Whole-Brain fMRI Results**
These images represent areas of increased activity from the whole-brain analyses displayed at a p threshold of .001, uncorrected.

Interactive Imagery > Implicit Baseline



Picture Imagery > Implicit Baseline



Picture Imagery > Interactive Imagery

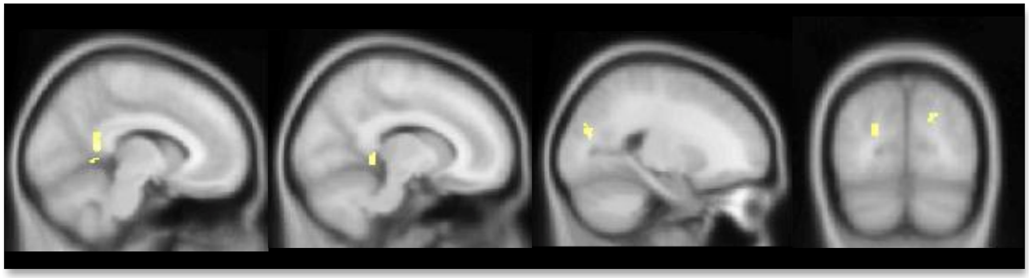


**Figure 8. fMRI Regression Results**

Figure (A) displays areas of increased activity related to larger learning slopes (greater increases in learning) for words learned through interactive imagery. Figure (B) displays areas of increased activity related to more negative retention slopes (greater loss of vocabulary information) for words learned through interactive imagery. Figure (C) displays areas of increased activity related to larger learning slopes (greater increases in learning) for words learned through picture imagery. Figure (D) displays areas of increased activity related to more negative retention slopes (greater loss of vocabulary information) for words learned through picture imagery.

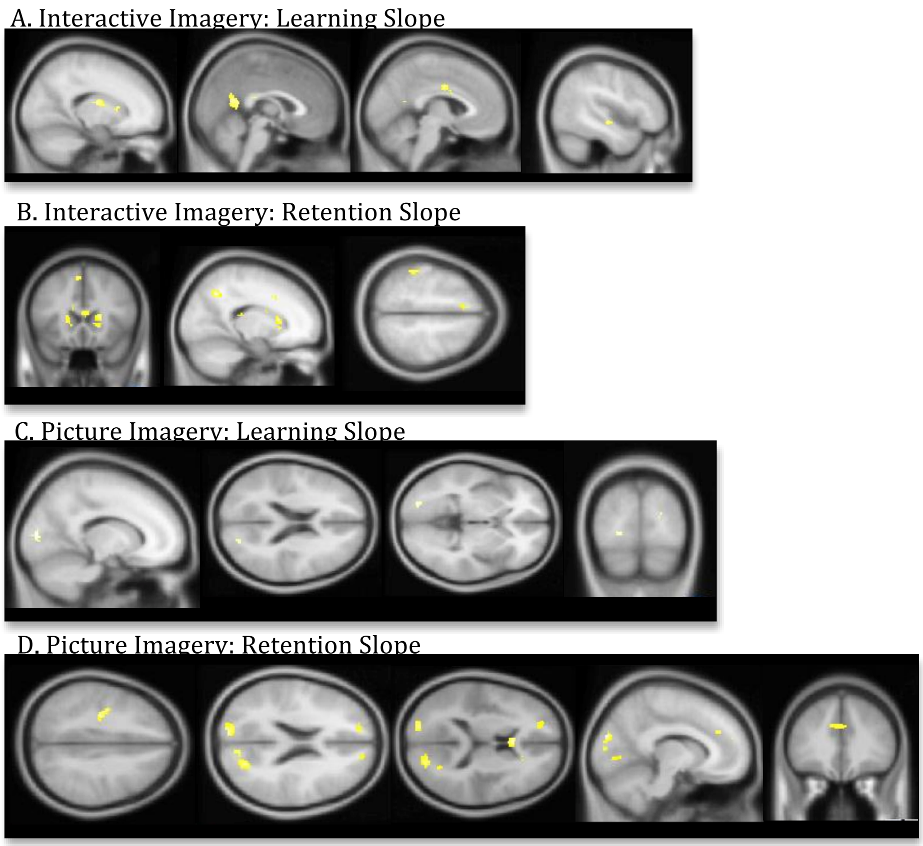
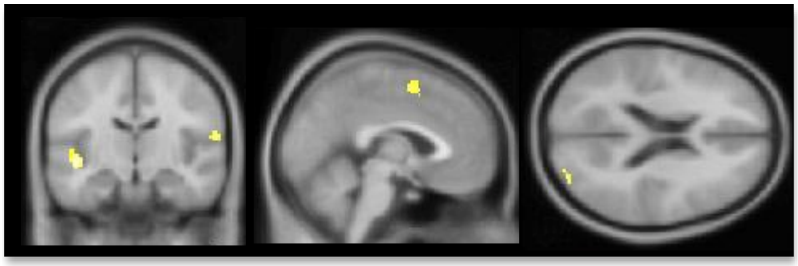


**Figure 9. fMRI Order Effects**

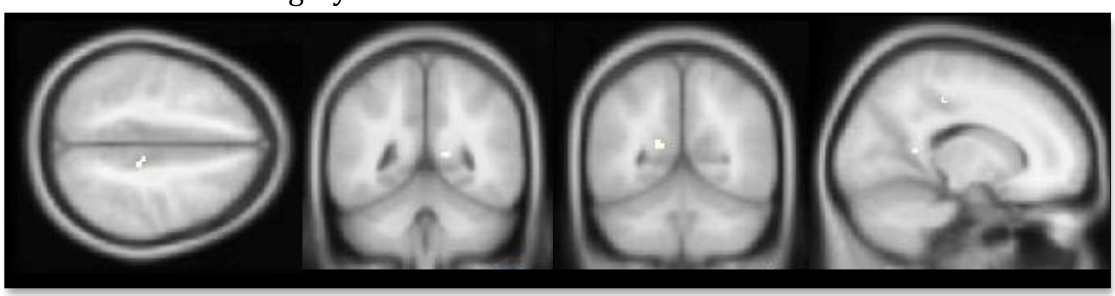
In order 1 (O1), interactive imagery training was given first. In order 2 (O2), picture imagery training was given first.

O1 > O2 Interactive Imagery



O1 > O2 Picture Imagery

O2 > O1 Picture Imagery



Appendix A: Training Words

| Group 1 Words | | Group 2 Words | |
|---|---|---|---|
| **English** | **German** | **English** | **German** |
| bag | tuete | cup | tasse |
| bandage | pflaster | key | schluessel |
| chair | stuhl | hose | schlauch |
| dog | hund | plate | teller |
| bottle | flasche | purse | tasche |
| pants | hose | fork | gabel |
| refrigerator | kiehlschrank | suitcase | koffer |
| pear | birne | glasses | brille |
| onion | zwiebel | coat | mantel |
| flower | blume | mirror | spiegel |
| boot | stiefel | asparagus | spargel |
| fly | fliege | pot | topf |
| light switch | schlater | button | knopf |
| picture | bild | knife | messer |
| cheese | kaese | teapot | kanne |
| piano | klavier | steering wheel | lenkrad |
| present | geschenk | slide | rutsche |
| pencil | bleistift | coin | muenze |
| clothespin | klammer | stove | herd |
| scissors | schere | whistle | pfeife |
| carrot | moehre | pliers | zanger |
| box | karton | cherry | kirsche |
| dustpan | schaufel | match | streichholz |
| belt | guertel | door | tuer |
| can | dose | sink | spuele |
| bicycle | fahrrad | window | fenster |
| bowl | schuessel | swing | schaukel |
| candle | kerze | sock | strumpf |
| cake | torte | lock | schloss |
| broom | besen | drum | trommel |