Bates, J. (2018). A theoretical framework for learning second/foreign language vocabulary in pictorial form. *Essays and Studies*, *68*(2), 139-159.

A Theoretical Basis for Learning Foreign Language Vocabulary in Pictorial Form

James William Bates

Introduction

The purpose of this essay is to examine a theoretical basis for the teaching of foreign language (FL) words in pictorial form. A theoretical bases of how language and visual information might be processed in the acquisition of learning foreign language vocabulary in pictorial form is proposed, consisting mainly upon three cognitive models: the dual coding theory (Paivio, 1986), the working memory model (Baddeley & Hitch, 1974), and Mayer and Moreno's (1998) cognitive theory of multimedia learning. The implications for FL vocabulary learning using pictorial representations of target words for each of these models is speculated upon, including what these models might mean for instructional situations, for both teachers and learners.

In FL and L2 (second language) learning situations, the vocabulary of the target language – along with translational forms – is sometimes presented to the learner as pictures in a variety of ways. A teacher might use flashcard featuring a picture of an object; using it to teach word meaning and then as a cue to word recall. Current L2/FL textbooks abound with images on practically every page, such as pictures and photographs of items and concepts relating to the text. Students may also use picture dictionaries, in which target vocabulary meaning is expressed as pictures. Modern vocabulary learning software (such as flash card applications) is multimodal, thereby having the capacity to support images of target vocabulary. Electronic glossing of text can also feature pictorial representations, where the pictorial expression of a word appears when highlighted. Although an emphasis exists upon understanding the meaning of words through translational means (especially within the Japanese

educational system) target words represented as pictures can still be a significant part of learning another language.

Learning and remembering FL vocabulary

When considering how words are learned, it is necessary to take into account the generally accepted theories that attempt to explain how people remember. Bartlett (1932) believed memory to be reconstructive as opposed to being reproductive: “An imaginative reconstruction or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience” (p. 213). Bartlett viewed memorization as an active process in which memories are constructed by an individual to their particular specifications, not simply a reproduction as with a video recording. A constructivist approach to memory regards the process as “the combined influence of the world, and the person’s own ideas and expectations” (Foster 2009, p. 13). Learning words in a foreign language can be described as a process involving the assimilation of form and meaning into an individual’s experience of the world. Apart from the frequency of encounters learners have of a word, the quality of the experience also influences their personal reconstruction of it.

The most accepted theories of vocabulary learning propose that words are not stored separately in the brain from other information, as information is so interconnected that it is impossible to identify lexical items as different entities (Hulstijn, 1997). The connectionist view of vocabulary knowledge does not see the mind as a kind of dictionary. Rather, it appears to rely upon the activation of connections, and connections within connections to store information (Hulstijn, 2001). Words are arranged systematically, in a large and highly complicated semantic network (Aitchison, 1987), and do not appear to be stored in isolation awaiting retrieval as with a dictionary entry. This complex arrangement is reflected in the idea that concepts (and thereby words) are categorized with flexibility in human memory, with Barsalou (1993) proposing that "Rather than being coherent, consistent, and complete, linguistic descriptions of conceptual content are unprincipled, haphazard, and incomplete" (p. 30). The connectionist view is apparent in the importance placed upon knowing the associations

(e.g., collocations, synonyms, words with opposite meanings, members of the same word family) of words to be learned. Nation (2013) believes that by knowing a word's associations, the word's "full meaning" can be better understood, and this knowledge "helps recall the word's form or meaning in the appropriate contexts" (p. 136).

Research into learning with pictures often refers to theories and models based upon a division between visual and linguistic information systems. Paivio's (1986) dual coding theory postulates two separate systems: verbal and nonverbal. Baddeley and Hitch's (1974) model of working memory features both audio processing involved with language and visual processing systems involved with imagery. Based upon these models, Mayer and Moreno's (1998) cognitive theory of multimedia learning describes how pictures and spoken words (as with a multimedia presentation) might be processed in the sensory and working memories in two separate channels.

Dual coding theory

The dual coding theory (Paivio, 1986) proposes that information is received simultaneously through two separate channels, one dealing with verbal and the other with non-verbal or visual information. The two types of input from verbal and non-verbal stimuli are received through the sensory systems, and are encoded and processed in their own separate way, as they have differing properties. Verbal information is language based and has a prepositional quality, so becomes a symbolic code with units referred to as logogens. Visual information being image based is more of a representation of actual physical objects in the real world, so becomes an analog code with units referred to as imagens. The two subsystems encode, organize, store and retrieve information differently and separately, yet this processing does include interconnections between the logogens and imagens, referred to as referential connections.

Using Paivio's (1986) dual coding theory as a theoretical basis, images may assist in the learning of FL vocabulary in the following way, as shown in Figure 2.1. A target word in pictorial (1) and FL written form (2) is received by the visual systems (3), and the

verbal form (4) is received by the auditory systems (5). Information from the picture is encoded as imagens in the nonverbal processing system (6), and FL written and verbal form are encoded as logogens in the verbal processing system (7). Processing involves referential connections between the two subsystems, so the inclusion of an image may assist learning as the nonverbal subsystem is more involved in processing than in the absence of an image.

The dual coding theory can be used to explain why images are more easily remembered than words, or the picture superiority effect. For example, when presented with a picture of a tree, the information (being an image) is channeled through the non-verbal subsystem as well as through the verbal subsystem as the linguistic form of *tree*. If receiving the word *tree* only, then the verbal (language related) channel will be active, but processing through the non-verbal channel will be minimal. As (according to the theory) semantic memory has a verbal and a visual encoding system, learning will be more effective if information is received in both verbal and visual modalities using a dual code, rather than if just one channel is used. If a word has an associated image stored along with a verbal entry, then there is a greater chance of retrieval.

The dual coding theory can also be used to explain why concrete words are better remembered than abstract words. According to the theory, concrete words (being easy to picture) are processed in both the verbal and visual subsystems. Abstract words, however, are only processed in the verbal subsystem (as they are difficult to visualize) so they stand a lesser chance of being remembered.

For FL vocabulary learning, the theory's basic premise that visual and verbal information is coded differently implies that pictorial representations of target words (in addition to verbal forms) should be presented to learners where possible. Adding another facet to the quality of the learner's experience of the target word is consistent with Craik and Lockhart’s (1972) levels of processing theory which explains how L2/FL vocabulary can be processed at different levels of semantic intensity postulating that the degree of semantic involvement with the word (the extent to which it is

considered, understood, and related to what is already known) is the main determining factor as to how the word is retained in the long-term memory.

Craik and Lockhart (1972) gave examples of three processing levels: (1) structural – a shallow encounter as with simply looking at the word in written form; (2) phonetic – a deeper encounter in which the sound form of the word is experienced; (3) semantic – regarded as "deep" processing, where the actual meaning of the word is considered. Providing a visual representation may facilitate 'deeper' processing of the target word than that achieved by experiencing textual and audio forms. In accordance with the dual coding theory, the codifying of words both visually and verbally provides the learner with a more meaningful experience than just focusing on the verbal channel only. Teachers should therefore provide visual representations of words where possible, so as to increase the chances of successful acquisition.

The working memory model

It would not seem possible (or even necessary) for humans to remember everything that they experience. Given the vast amount of visual and acoustic information alone which is received on a daily basis, the need for the brain to select what is to be committed to memory is apparent. Based on empirical evidence, it has been established that a kind of temporary store does exist, where seven (plus or minus two) items can be processed (Miller, 1956) for a duration of 15 to 30 seconds (Atkinson & Shiffrin, 1971). A model of such a compartmentalized memory system is the multistore model (Atkinson & Shiffrin, 1968), which proposes three separate memory storage systems. According to the model, input from the environment is received by the sensory memory, and then information which has been subject to attention is received by the short-term memory. Information is held in the short-term memory, where some of it is "rehearsed", resulting in transference to the long-term memory.

The working memory model (Baddeley & Hitch, 1974) offered a more complex model of memory than did the multistore model, with the concept of a short-term memory being replaced by the more complicated working memory. According to the model, as

with the dual coding theory, information is received in an auditory and a visual channel. The processing of auditory information occurs in the working memory's phonological loop, and visual information is processed (represented and manipulated) in the working memory's visuo-spatial sketchpad. These two sub-systems are managed by a separate control system, the central executive. Processing occurs within and between each sub-system, before information is stored in the long-term memory. Baddeley (2000) later added the episodic buffer to the model, which works as a kind of back-up or reserve storage system, and participates in the communication of information between the central executive and the long-term memory.

In accordance with the working memory model, pictorial expressions of target words might influence the learning of FL vocabulary in the following way, as shown in Figure 2.2. The image (1), the FL written form (2), and the FL verbal form (3) are received as input by the sensory memory (4). Information attended to is received by the working memory, where the image and the FL written form are processed in the visuo-spatial sketchpad (5), and the verbal form is processed in the phonological loop (6). Information and processing is managed by the central executive (7), from where target word information can be transferred to the long-term memory (8).

Baddeley (2003) described the working memory's visuo-spatial sketchpad subsystem as serving "the function of integrating special, visual, and possibly kinesthetic information into a unified representation which may be temporarily stored and manipulated" (p. 200). An important quality of the visuo-spatial sketchpad is its flexibility; its ability to manipulate mental images such as, for example, when an engineer might "turn around" a building plan in their mind to answer a specific question. Images are not stored as static entities, as with photographs and videotape. Rather, they are more like image and video files which can be altered at will with a kind of editing software. Baddeley (2004) stated his belief that images cannot be directly stored in the brain, as the sheer amount of information an image has would be far too demanding on its storage system, so images in the long-term memory are likely to be stored in a kind of abstract code. However, the working memory may use a system that is more of a direct representation

of the image, which Baddeley (2004) describes as a "spatial slave system" which uses "the same equipment as used in perception, and depends for its functioning on the central executive component of the working memory" (p. 59).

As evidence in support of this view, Baddeley (2004) referred to an experiment by Brooks (1968) in which participants "rotated" a capital letter F in their minds, while answering simple questions about how the letter appeared in their minds. Participants in his study found that responding to questions related to the image was more difficult when answering spatially (when pointing to answers on a paper) than verbally (when giving an oral response). It appears that in order to give a response that requires visual and spatial processing; the limited visual and spatial resources in the working memory are diverted, leaving it with less capacity to carry out the visual/spatial task. Yet this was not the case with verbal resources, as interference to the task from verbal responses was significantly less. Brooks (1968) achieved similar results when his participants performed a task requiring verbal processing, where verbal responses were inhibited more than visual/spatial responses. These experiments suggest (in addition to the existence of verbal and visual subsystems of limited capacity) that visual information encoded in the working memory might be processed visually and spatially. Although they are not like photographs inside the brain, mental images may still have a kind of representational quality, which allows for manipulation in order to facilitate understanding and memorization.

Baddeley (2004) believed that imagery may also play a significant role in the storage of information in the long-term memory, as evidence exists that the capacity of a word to be imagined can determine how well it will be remembered, and the fact that imagery plays a prominent role in the use of mnemonic strategies. Based on earlier experimentation (Baddeley, Grant, Wright, & Thompson, 1973), he concluded that mnemonic techniques involving imagery rely upon the visuo-spatial sketchpad, as a visual/spatial task interfered with the use of a mnemonic, with no difference between the remembering of concrete or abstract words. In another experiment, no significant difference in the interference from a visual/spatial task was found between the

remembering of concrete and abstract noun-adjective word pairs, indicating that the concreteness of a word is not mediated by the visuo-spatial sketchpad. Baddeley (2004) said, “The concrete/abstract difference provably has something to do with the way the word characteristics are stored in semantic memory, with concrete words being more richly encoded than abstract ones” (p. 62).

More recently, Baddeley (2015) has elaborated upon his earlier models with the inclusion of colour, shape, spatial location, and kinaesthetic elements into the concept of the visuo-spatial sketchpad, and the inclusion of non-audio communication (such as sign language and lip reading) into the phonological loop. With regards to L2 language learning, Baddeley (2015) maintains that a proven link exists between the phonological loop and the long-term memory’s capacity to learn language, stating: “Not only does the capacity of the phonological loop influence the rate of vocabulary acquisition, but also, conversely, a richer vocabulary is associated with increased verbal memory capacity” (p. 24). However, with regards to the visuo-spatial sketchpad, he says that research has been scant in this area, and speculates that the visual subsystem may play a role in the acquisition of orthographic systems which are visually complex as with Chinese characters.

Baddeley’s various models have been challenged by other models of working memory, which are often process orientated rather than structurally orientated (Bunting & Engle, 2015). For example, Cowan’s (2005) embedded process model theorizes that a capacity limited focus of attention plays a significant role in memory processing. However, Cowan (2015) acknowledges the important role of phonological and visual processes working within the working memory, having stating that retention of L2 vocabulary in the long term memory requires phonological forms and “A few recent visual events that may be candidates for the meaning of the new word” (p. 32). According to Wen (2016) the general consensus amongst cognitive psychologists is that there are domain specific (i.e., visual and verbal) mechanisms within the working memory, as “a completely unitary, domain-general view of working memory does not hold” (p. 21).

How the working memory functions and how it can actually be defined remains speculative. However, Baddeley's original view of working memory remains highly influential. According to Wen (2016), the model's apparently simplistic three part framework "has proven to be an extremely powerful framework for addressing a range of questions on high-level human activities" (p. 13). With regards to L2/FL acquisition, research into the working memory remains extensive, especially with studies concerning the phonological subsystem (Baddeley, 2015). Working memory is also currently believed to be a fundamental and highly influential factor upon language aptitude (Ellis & Shintani, 2014).

The working memory model suggests that when experiencing pictures in order to learn FL vocabulary, students need sufficient time to process what they are learning. If a construction of meaning is actually taking place within the learner -in which visual and verbal elements are being processed in order to build an adequate 're-presentation' of target word meaning to be coupled with its verbal and written form- allowing students sufficient time to manage and manipulate their constructions may be crucial to learning. The model's visual processing subsystem (the visuo-spatial sketchpad) alone appears complex and demanding. It is not perceived to be an instantaneous process (analogous to a photographic plate or a computer scanner) but more like a blank sketchpad requiring action on the learner's part. Like any drawing, students will need time and effort to do their mental sketching. Time is also required for other processes such those involved with linking the target word concept with its verbal form being processed by the phonological loop, as well as the building of associations with existing knowledge and integration with the learners' long term memory.

The model also highlights the concept of memorization as being an active process. For FL vocabulary study, this quality of human memory implies that learning with pictorial information should be an active rather than a passive endeavour. An active style of learning might be encouraged by explicit and deliberate engagement with the pictorial representations of target words. For example, images presented in class could be examined, discussed, and evaluated. A pictorial representation of a word often lacks the

accuracy to convey meaning, especially when representing less concrete and more abstract concepts. For example, the word *car* will always be easier to represent than the word *legislate*. Rather than simply accept an image as being an accurate portrayal of its referent, students should be encouraged to scrutinize and criticise the chosen picture: Is it an accurate representation? What are its strengths and weaknesses as a pictorial expression of the target word? Students could also be involved in the selection process of pictures, or in the actual creation of them. Such an explicit approach to learning with images may result in greater retention of target words, as according to the working memory model, human memory is not simply a passive storage vessel waiting to be filled. The model proposes that, when learning FL vocabulary with pictures, focused attention and a concerted cognitive effort are required in order for the learner to construct their own internal (mental)representations of target word meanings and concepts from the external representations (the pictures) presented.

Cognitive theory of multimedia learning

Mayer and Moreno's (1998) cognitive theory of multimedia learning offers another theoretical model of how words and images are received, processed and stored. The theory is primarily concerned with how different modes of information should be presented to the learner in order to maximize the learning experience. Having a practical quality, it is often referred to in literature concerning the design and use of educational and presentation material in computer-assisted language learning (CALL) environments, as well as studies concerning the use of pictorial information in instructional settings. The theory is based upon other theories and models as seen in the theory's features, such as the two information channels of the dual coding theory (Paivio, 1986), the multi-store memory systems of the working memory model (Baddeley & Hitch, 1974), and the restricted processing capacity of Sweller's (1988) cognitive load theory.

The theory is based upon three main assumptions concerning human cognitive processing; "The human mind is a dual-channel, limited-capacity, active-processing system" (Mayer, 2005, p. 37). Firstly, the system is dual-channel, as information follows two specific pathways, originating in the multimedia presentation and ending up

in the long-term memory. These channels transform the information from five different sources: (1) the words and pictures of the presentation itself; (2) the acoustic representations (sounds) and iconic representations (images) in the sensory memory; (3) sound and images in the working memory; (4) verbal and pictorial models also in the working memory; and (5) schemas which are stored in the long-term memory, becoming part of existing knowledge (Mayer, 2005). Secondly, the process has a limited-capacity as the two subsystems working simultaneously can only process a finite amount of information at one time, similar to that of the limited processing power of a computer. Thirdly, the process is an active-processing system. The words and pictures of a multimedia presentation are received in the sensory memory through the ears and eyes. Then, according to Mayer (2005), information is actively processed in the working memory in three different ways: (1) selecting sounds (the words) and images from the sensory memory; (2) organizing the sounds and words into verbal and pictorial models; and (3) integrating these models into the prior knowledge of the long-term memory. The learner is not a passive agent when receiving multimedia material.

As an active participant, the learner is charged with understanding the material presented and then constructing mental models that will eventually become part of their existing knowledge. Multimedia presentations need to support and encourage the learner in this process, as "One of the principle aims of multimedia instruction is to encourage the learner to build a coherent mental representation from the presented material" (Sorden, 2012, p. 2). The use of visual information is central to this process as it is theorized that successful learning of a multimedia presentation requires the formation of meaningful links between words and images. Meyer (2001) regards the construction of connections between word-based and image-based representations as the most critical stage in multimedia learning. As with dual coding theory, the addition of both visual and verbal input will result in a more effective learning experience than with only one form. With Mayer and Moreno's theory, however, more emphasis is placed upon verbal and visual processes working together to build internal representations of what has been presented to the learner.

According to Mayer (2009), the three processes of selection, organization, and integration result in meaningful learning, as opposed to no learning or shallow rote learning, as with the previously mentioned levels of processing theory (Craik & Lockhart, 1972) which postulated that the greater the degree of thought put into what is being attended to, then the better the chances of learning. In the case of Mayer and Moreno's theory, meaningful learning is regarded as selecting relevant words and images for the working memory, organizing the selected words and images into verbal and pictorial models, and then integrating these models with each other and with prior knowledge.

As shown in Figure 2.3, the presentation of pictorial representations of target words (in accordance with the cognitive theory of multimedia learning) may assist the process of learning English vocabulary in the following way. The multimedia presentation consists of the target word in written FL form (1) and verbal FL form (2) presented as words, and the presentation of pictorial forms (3). The senses (4) receive the FL written form and the pictorial form through the eyes, and FL verbal form through the ears. The working memory (5) selects and organizes the words and pictures, creating verbal and pictorial models. Some target words are integrated with prior knowledge into the long-term memory (6).

In order to promote meaningful learning in multimedia presentations, cognitive processes need to be supported. Meyer, Fennell, Farmer, and Campbell (2004) identified two important considerations. Firstly, that cognitive load should be reduced so that the working memory (which has a limited capacity) is freed up or made available to carry out the three processes. Secondly, learners should be encouraged to use this available cognitive "space" by providing material which catches their interest, thereby encouraging a deeper level of processing.

As with other models of cognition, such as the dual coding theory, cognitive processes are viewed as having a limited capacity. Sweller's (1988) cognitive load theory argued that instruction should be designed so as to reduce the learner's cognitive load. In line

with Sweller's theory, Mayer (2009) specified twelve multimedia instructional principles, designed to support the learner in their multimedia experience by controlling and economizing their cognitive task. Mayer grouped the principles in accordance with Sweller's (1988) classification of cognitive load, as supporting extraneous, essential or generative processing.

Extraneous processing refers to the processing of information that is unnecessary to what is being learned. Mayer (2009) believes that extraneous processing results in "cognitive processing that does not serve the instructional goal" (p. 57), so information superfluous to what is being taught needs to be reduced, as it hinders learning by taking up limited cognitive processing resources. Meyer recommended five principles of multimedia instruction which he believes would reduce extraneous processing: (1) Coherence Principle – Excluding rather than including extraneous material; (2) Signaling Principle – Providing cues that highlight the organization of the essential material being added; (3) Redundancy Principle – Presenting graphics and narration rather than graphics, narration, and printed text; (4) Spatial Contiguity Principle – Placing corresponding words and pictures near each other; (5) Temporal Contiguity Principle – Presenting corresponding words and pictures at the same time rather than in succession.

Essential processing refers to the processing of information necessary for the material presented to be understood. Mayer (2009) describes essential processing as "cognitive processing that is required to represent the material in working memory and is determined by the complexity of the material" (p. 57). Information needs to be presented in a way that does not exceed the learner's ability to receive and process it. If the working memory cannot process information properly then learning will not occur, so essential information needs to be managed. Mayer (2011) identifies three principles which encourage the control of essential information: (1) Segmenting principle – Information presented in user-paced segments rather than as a continuous unit; (2) Pre-training principle – Giving the names and characteristics of key components in

advance; (3) Modality principle – Presenting graphics and narration rather than graphics and printed text.

Generative processing is processing activity aimed at developing a deeper understanding of the material through "organizing the incoming material into coherent structures and integrating these structures with each other in prior knowledge" (Mayer, 2009, p. 221). Principles supporting generative processing focus on the audio and visual modes in which information is received: (1) Multimedia principle – Words and pictures are better than words alone; (2) Personalization principle – Words in a conversational rather than in formal style; (3) Voice principle – A friendly, human voice rather than a machine voice; (4) Image principle – The speaker's image is not necessary when the voice is being heard.

Mayer has developed and modified his list of principles based on experimental research (see Mayer, 2011), yet the principles continue to follow the basic concepts of extraneous, essential, and generative cognitive processing. In simple terms, Mayer suggests that, in accordance with his model, multimedia presentations should be concise, well-organized, and thought provoking. Mayer and Moreno's (1998) cognitive theory of multimedia learning appears to provide a useful guide for the development of instructional materials, based on a theory of how learners learn. Sorden (2013) believes that the theory is relevant to current educational needs, as it is "learner-centred and has a cognitive constructivist orientation" (p. 168). He further states that the cognitive theory of multimedia learning will continue to evolve and to be relevant as it "focuses on finding effective instructional methods" (p. 168), so therefore is not bound to any instructional methods which can become redundant.

The cognitive theory of multimedia learning has some important practical implications for the teaching of FL vocabulary with pictures. The model emphasises the limited capacity of learners to engage in meaningful learning, as the processes of selection, organization and integration of information are all dependent upon finite cognitive resources. From the model's perspective, the FL vocabulary learner experiencing FL

vocabulary pictorial form has several cognitive 'tasks' to perform, including the complex interaction and construction of meaning between verbal and visual representations of target words. Alleviating the cognitive load on these processes, through the limiting, discouragement, and removal of extraneous factors to learning, appears necessary in order to increase the chances of successful target word acquisition.

Put simply, the learner should not be overloaded with information. Pictures should be presented one at a time, accompanied by either the verbal form of the target word or the written form, without the inclusion of any superfluous material. Word cards (made of paper-based medium) can provide such a system, where word forms are presented as discrete, single units; at a pace which can be controlled by the individual learner. According to Nation (2013) electronic flashcards can also be an effective means of learning L2 vocabulary. Electronic flashcards -now appearing as smartphone applications- have many advantages over traditional paper cards, such as the capacity to automatically employ learning systems (such as the Leitner system) designed to maximize the conditions for successful memorization. Electronic flashcards also have a significant advantage in their capacity to support multimedia, as target FL vocabulary can be experienced in written, pictorial, verbal, or even video form. Images can be selected and created by the learners themselves, and these can be easily shared amongst users.

Cognitive load might also be alleviated by the visual qualities of the images themselves. In order to reduce the amount of visual input the learner experiences, images with a simplistic representation of their referent word might be preferable, as with the examples shown in Figure 4. Pictures could consist of simple line drawing instead of more detailed designs, and black and white used instead of complex shading and colouring. Road signs have this pictographic quality, being designed for the observer to quickly and easily receive the sign's meaning, thereby instantly understanding the intended referent (e.g., a silhouette of a deer means "watch out for deer"). In a similar way, simple images might reduce the cognitive load upon the learner, thereby supporting their cognitive processing of word form and meaning, and eventual

integration of the new word into the long term memory. Additionally, the simplistic forms (lacking detail) may encourage students to construct their own internal representations of meaning, as the simple image requires students to use their own knowledge to construct an internal representation of meaning. This process may result in a deeper and more personalized understanding of the FL words.

Conclusion

This essay has examined a possible theoretical basis for the teaching and learning of FL vocabulary with vocabulary items represented pictorially. Human memory is not like a recording but is constructive and involves existing knowledge, with words being stored not as isolated entries but within complex networks of meaning. The dual coding theory (Paivio, 1986) emphasises the importance of presenting information visually as well as verbally, as using two channels will result increase the semantic involvement of the learner with the target vocabulary more so than if only one channel is used. The working memory model (Baddeley & Hitch, 1974) proposes that when given a pictorial representation, learners construct an internal representation of word meaning in a temporary store by creating meaning from the encoding and interaction of visual and verbal information. This process suggests that learners require adequate time to allow these events to occur, with FL vocabulary learning being more of an active than a passive process requiring explicit engagement and focused attention. Mayer and Moreno's (1998) cognitive theory of multimedia learning emphasises the limited capacity of human cognition, suggesting that the cognitive load upon learners needs to be reduced in order for successful learning to take place. For FL learning with pictures, this could mean that images should be presented so as not to overload the learner with information, and the images themselves should be simplistic.

All three models differ yet have a fundamental quality of having visual and non-visual (verbal, language-related) subsystems, emphasising the important role that the use of imagery in FL vocabulary learning might play. Students could be given a visual representation of target words in order to support them in their effort to construct their own internal representations, helping them to be able to retain and recall the words

successfully. Recent innovations such as electronic flashcards used in mobile technology may provide excellent opportunities for students to study FL words in not just as L1 translational, FL verbal, or textural forms, but in pictorial form which may be a positive and effective influence upon their learning.

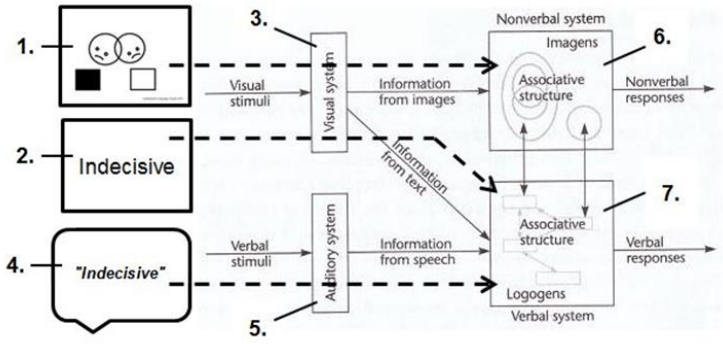

Figure 1. Paivio's (1986) dual coding theory as a theoretical basis to learning FL vocabulary with pictorial representations. Diagram adapted from Ware (2004).

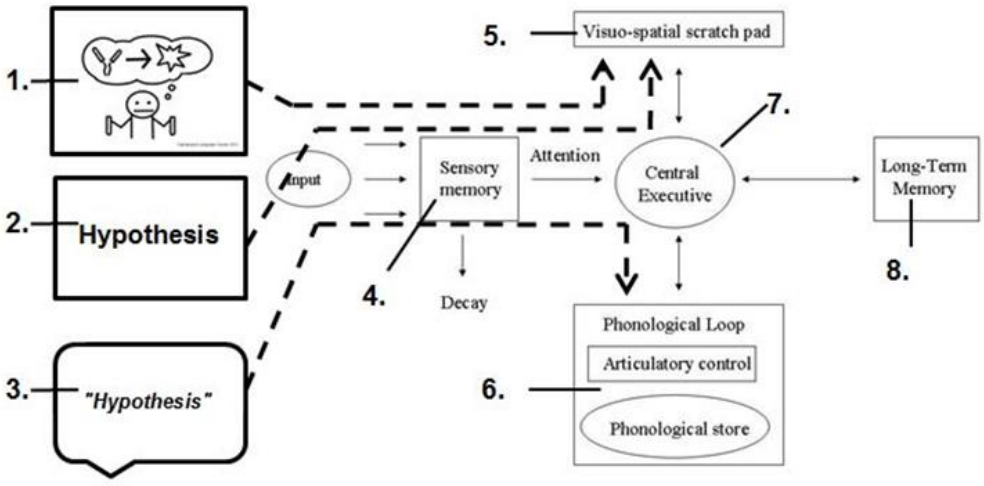

Figure 2. Baddeley and Hitch's (1974) working memory model as a theoretical basis to learning FL vocabulary in pictorial form. Diagram adapted from McLeod (2012).

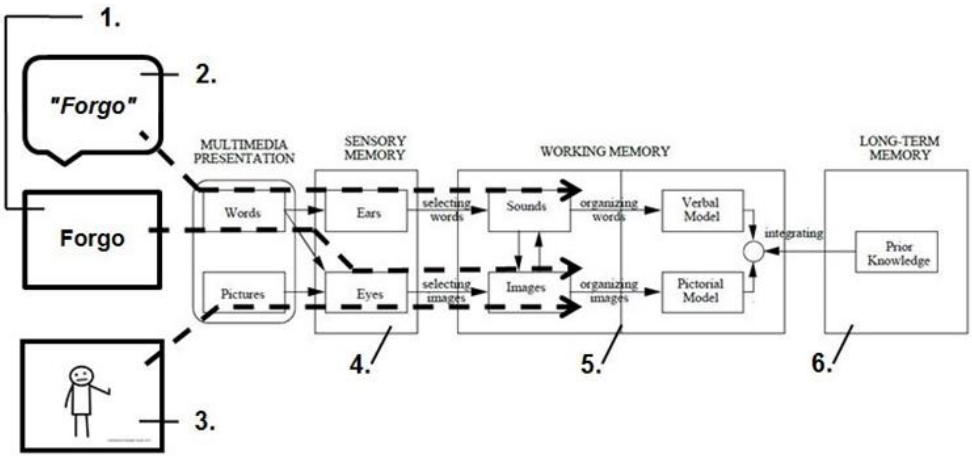

Figure 3. Mayer and Moreno's (1998) cognitive theory of multimedia learning as a theoretical basis to learning FL vocabulary with pictorial representations. Diagram adapted from Mayer (2011).

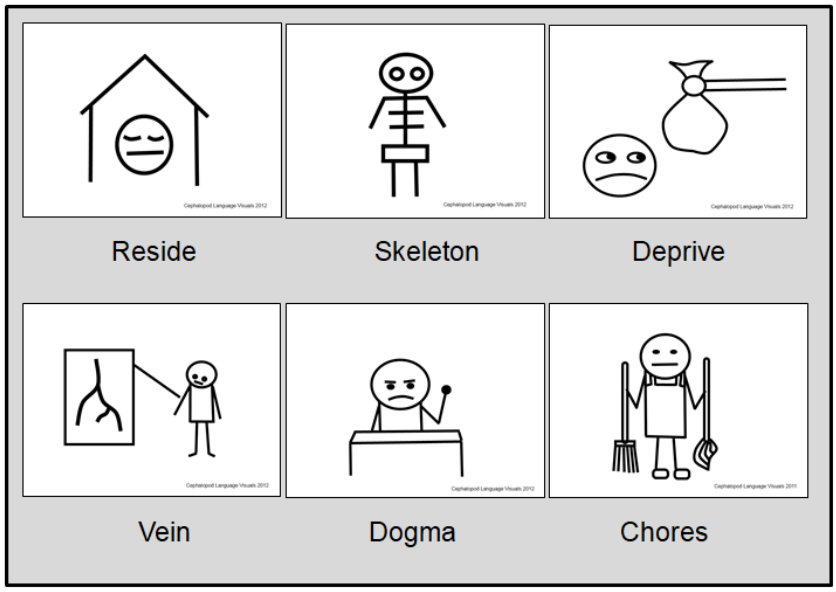

Figure 4. Example simple pictorial representations of English words. Pictures designed by the author.

References

Aitchison, J. (1987). *Words in the mind.* Oxford: Basil Blackwell.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R.M. (1971). The control of short term memory. *Scientific American*, *225*(2), 82-90.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417-423.

Baddeley, A. D. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*, 189–208.

Baddeley, A. D. (2004). *Your memory: A user's guide* (3rd ed.). UK: Carlton Books.

Baddeley, A. D. (2015). Working memory in second language learning. In Z. Wen, M. B. Mota & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 17-28). UK: Multilingual Matters.

Baddeley, A. D., Grant, S., Wight, E., & Thomson, N. (1973). Imagery and visual working memory. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 205-217). London: Academic Press.

Baddeley, A., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation*, (pp. 47-89). New York: Academic Press.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. Collins, S. Gathercole, M. Conway & P. Morris (Eds.), *Theories of memory* (pp. 29-101). UK: Lawrence Erlbaum Associates, Ltd.

Bartlett, F. (1932). *Remembering.* Cambridge: Cambridge University Press.

Brooks, L. R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology*, *22*, 349-368.

Bunting, M., & Engle, R. (2015). Foreword. In Z. Wen, M. B. Mota & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. xvii-xxiv). UK: Multilingual Matters.

Cowan, N. (2005). *Working memory capacity*. Hove: Psychology Press.

Cowan, N. (2015) Second language use, theories of working memory and the Vennian mind. In Z. Wen, M. B. Mota & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 29-40). UK: Multilingual Matters.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.

Foster, J. K. (2009). *Memory: A very short introduction.* Oxford: Oxford University Press.

Hulstijn, J. (1997). Mnemonic methods in foreign language vocabulary learning. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp.174-200). Cambridge: Cambridge University Press.

Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.

Mayer, R. E., Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. Retrieved from https://gustavus.edu/education/courses/edu241/mmtheory.pdf

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E. (2011). Instruction based on visualizations. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 427-445). New York: Routledge.

Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, *96*(2), 389-395.

McLeod, S. A. (2012). Working memory. Retrieved from www.simplypsychology.org/working%20memory.html

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on the capacity for processing information. *Psychological Review*, *63*, 81-97.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.

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

Sorden, S.D. (2012). Cognitive Theory of multimedia learning. Retrieved from http://sorden.com/portfolio/sorden_draft_multimedia2012.pdf

Sorden, S.D. (2013). Cognitive theory of multimedia learning. In B. J. Irby, G. Brown & R. Lara-Alecio (Eds.), *Handbook of educational theories* (pp. 155-168). Charlotte: Information Age Publishing Inc.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Ware, C. (2004). Information Visualization: Perception for Design (2nd ed.). San Francisco: Morgan Kaufmann Publishers.

Wen, Z. (2016). *Working memory and second language learning.* UK: Multilingual Matters.