Second-Language Vocabulary Learning and the Additivity Hypothesis

François Pichette
University of South Florida

Paivio's (1975) additivity hypothesis suggests that simultaneous stimuli from different sources enhance the recall of an item. However, little is known about the additivity of pictures and spoken words in reading conditions where participants are allowed time to generate inner speech and mental imagery. The present experiment investigates the existence of additivity in the learning of second-language concrete vocabulary by presenting up to three types of stimuli simultaneously. Participants are 50 anglophone college students learning Spanish as a second language. The results obtained suggest that when inner speech and mental imagery occur, adding extra stimuli does not improve recall. Under these circumstances, recall seems to depend more on word length and on individual learning strategies that probably consist of mental rehearsing and linking new items to already stored information.

Selon l'hypothèse de l'additivité de Paivio (1975) des stimuli de différentes sources favorisent le rappel d'un item. Par contre, on sait peu de choses sur l'additivité d'images ou de mots prononcés dans des conditions de lecture où les sujets ont le temps de se créer des images mentales et de la parole interne. La présente étude se penche sur l'existence de l'additivité pour l'apprentissage du vocabulaire concret en langue seconde en présentant simultanément jusqu'à trois types de stimuli. Les sujets sont 50 étudiants anglophones de niveau universitaire apprenant l'espagnol comme langue seconde. Les résultats obtenus suggèrent que quand le langage intérieur et l'imagerie mentale surviennent, l'ajout de stimuli supplémentaires n'améliore pas le rappel. Dans ces circonstances, le rappel semble dépendre davantage de la longueur du mot et des stratégies d'apprentissage individuelles qui consistent probablement en des répétitions mentales et en l'association des nouveaux items aux connaissances déjà acquises.

Introduction

Over the last decades, studies have consistently shown pictures to be recalled better than words across several types of memory tasks (see Madigan, 1983 and Mintzer and Snodgrass, 1999). For example, Paivio and Csapo (1973) compared the efficiency of pictures versus words for recall, and their results

Address for correspondence: François Pichette, Department of World Language Education, University of South Florida, 4202 East Fowler Ave., CPR 419, Tampa, FL 33620-5550, USA. E-mail: pichette@chuma1.cas.usf.edu.

suggest that one occurrence of a picture is generally worth two of a word. One of the explanations that has received the most attention for the picture superiority effect is Paivio's dual coding hypothesis, according to which pictures are better remembered because they are more likely to be encoded both as images and as verbal traces, while words elicit just a single verbal code (Paivio, 1971). This possibility of dual coding is also suggested by Baddeley's (1986) model of working memory, which contains a phonological loop and a visual-spatial sketch pad as parallel systems.[1] Paivio (1975) formulated the additivity hypothesis, which states that a word accompanied by a picture of the concept shows an additive effect on recall when compared to the recall of a word or picture alone. These hypotheses also tend to account for the superiority of concrete words over abstract words (Paivio, 1965; Paivio, Walsh and Bons, 1994; Richardson, 1980, for example), the former triggering mental images that would favour better recall.

In addition to triggering mental images, written words also have a phonological equivalent that the reader can generate through inner speech, which is seen as a normal part of silent reading and as being automatically generated (Underwood and Batt, 1996; Zhang, Perfetti and Yong, 1999). Inner speech manifests itself by covert movement of speech muscles during verbal thought called *subvocalization*. Even though subvocalization increases with textual difficulty (McGuigan, 1970), it is less clear whether inner speech serves a useful purpose. Numerous researchers argue that the phonological form of written words helps to facilitate lexical access and reading comprehension, either for all words or as a compensatory strategy for words that are not recognized automatically, through phonological mediation (Lee, Kim, Binder, Pollatsek and Rayner, 1999). Other researchers argue that phonology can help hold word information in working memory (Gathercole and Baddeley, 1990), or that information about prosodic structure may remain available and help retrieval of the concept (Slowiacek and Clifton, 1980). Hence, the phonological form of a word, be it heard or self-generated, is additional information that can guide the reader in identifying and understanding written material, as suggested by the additivity hypothesis.

Despite their high number, earlier studies that compared verbal and nonverbal stimuli have not provided language instructors with much information about the usefulness of diversifying the stimuli for enhancing learning and recall of new words in a second language. The first reason is that most studies are based on Stroop tasks (Stroop, 1935) and are interested in the phenomenon of interference on categorizing or naming tasks. For example, colour names are written with an ink colour different from the name being written (Dyer, 1973) or distractor words accompany pictures (Rosinski, Golinkoff and Kukish, 1975; Lupker, 1985). The most common goal is to examine word identification by focusing on the process instead of on effects on learning and recall. In addition,

the majority of studies have been carried out in L1 with concrete and frequent items, implying that very few items are unknown and learned. Another reason earlier studies have not provided answers for language instructors is that simultaneous use of different types of stimuli for the same item is rare, and comparisons with the absence of the added stimuli are even more scarce, thus saying nothing about additivity.[2] Finally, the rate of presentation in the studies is usually much faster than in normal teaching conditions (for example, 0.12 sec. in Paivio and Csapo, 1973), and less than the three seconds judged necessary for image generation to occur (Paivio, 1966). Under these circumstances, most previous studies did not consider imagery or inner speech, which are processes involved in L2 vocabulary learning. The few studies which actually investigated imagery and additivity of stimuli (Gildea, Miller and Wurtenberg, 1990; Mousavi, Low and Sweller, 1995; Tindall-Ford, Chandler and Sweller, 1997, for example) were for the most part conducted on sentences or text instead of isolated words and focused on content learning in the L1.

Given these facts, it is clear that studies in this area of investigation, although numerous, provided results that were in most part not applicable to language teaching and learning. Tasks used in the psychology laboratory that consist of naming an item, recognizing it, categorizing it or pairing items do not involve the same processes as those involved in vocabulary learning. The presentation of the items does not usually reflect the way new words are normally introduced in language courses: teachers seldom teach words in semantically unrelated pairs, nor do they introduce several new items in less than one second or resort to interference as a teaching method. In typical L2 classroom practice the new written word does not disappear from sight, and the learner has time to generate mental imagery, produce inner speech, rehearse the word mentally and reflect on it.

The present study investigates the existence of additivity in the learning of concrete words in the L2, using simultaneously up to three different types of external (not self-generated) stimuli in two sensory modalities. According to the additivity hypothesis, adding stimuli from other sources in the presentation of an item will result in better recall than presenting the written word only. However, given the considerations previously mentioned, it is not certain whether the addition of the spoken form will have a significant additive effect on the verbal trace created by inner speech, nor is it certain that adding a picture will have any additive effect for participants who had been given time to create a mental image.

It is difficult to predict whether high-imagery words might be recalled equally well with or without accompanying pictures. This question is at the root of the long debate about the pictorial nature of mental images (Pylyshyn, 1973; Anderson, 1978). On the one hand, imagery and perception share common mechanisms and activate common brain areas (Trojano and Grossi, 1994;

Kosslyn, Thompson and Alpert, 1997; Mellet, Petit, Mazoyer, Denis and Tzourio, 1998). On the other hand, there are individual differences in the fluency and vividness of mental imagery (Marschank and Cornoldi, 1991). Further evidence also suggests that imagery and perception have different properties: while imaging seems to be bound by a prototypical frame of reference for an object, perception involves an independent object that can redefine this original frame and allow for closer examination (Kaufmann, 1996; Barsalou, 1999). Similar doubts about additivity apply to auditory stimuli when added to inner speech: would hearing a word help recall if that same word in the written form were presented long enough for the participant to create a phonological equivalent? The auditory stimulus could have an advantage over inner speech in the case of new words, since it carries pronunciation and prosodic information previously unknown to the participant. It is doubtful that new phonological information could bring about relevant images, as do written words that are already known in their spoken form. Also, the phonology a reader associates with a written word might differ from the real phonological equivalent.

Another variable to consider in the design is word length. Most studies on word learning from lists show poorer recall for longer words. The dominant explanation is that items which take a longer time to articulate allow for less rehearsal, and less opportunity for learning associations between positions and items (Baddeley, Thompson and Buchanan, 1975). Another explanation relates the effect to decay during output (Cowan et al., 1992; Avons, Wright and Pammer, 1994). Word length is expected in this experiment to correlate negatively with recall.

The present study addresses the following questions for participants reading lists of isolated words when they are given time to generate mental images and inner speech:

– Is recall enhanced when a picture representing the concept is provided?
– Is recall enhanced when the word is also heard by the participant?
– Is recall influenced by word length?

Method

This section will provide information on the individuals who participated in the experiment, the materials that were used and the analyses that were performed on the data.

Participants

Participants were 50 university students from a diversity of majors who were enrolled in introductory Spanish courses at the University of Akron, Ohio. The participants were 23 males and 27 females, ranging in age from 17 to 52 (average 22.4).

Materials and procedure

Stimuli consisted of names of animals and food items (mainly fruits and vegetables). These items were selected because they are part of the course content, and because they pertain to semantic fields in which words should have a high capacity for imagery. Word length varied from two to four syllables and from four to eleven letters. Frequency was a selection criterion since participants have to know what each word refers to in order to create a mental image without an accompanying picture. Less common words such as *endive* could bias the recall process by not allowing some participants to create a mental representation. Cognates were also eliminated to avoid the addition of a crosslinguistic variable, since the participants would otherwise not be dealing with entirely "new" words. Also, it is commonly assumed that new words are incorporated through the detection and use of similarities between new and already stored information (Matz, Teschmer and Weise, 1988)

Three series of 12 items were prepared, each having an equal number of food and animal items. Presentation was made using the Microsoft Power Point and the Sony Conductor Companion software. Each series consisted of three items for each of the following conditions, A through D:

A: Spanish word + English equivalent

B: Spanish word + English equivalent + pronunciation

C: Spanish word + English equivalent + picture

D: Spanish word + English equivalent + picture + pronunciation

For Conditions A and B, words were on the same line in the centre of the screen, with the Spanish word (in font Times New Roman size 60) followed by the English equivalent in parentheses (in font Times New Roman size 44). For Conditions C and D, the same written configuration was in the top half of the screen and the coloured picture was in the bottom half. All items were centred horizontally. Each picture represented one typical item. Figure 1 shows slide samples for Condition A (*ardilla* 'squirrel') and Condition C (*ballena* 'whale'). Adding the pronunciation to these samples would result in, respectively, Condition B and Condition D.

A supplementary item was presented at the beginning and at the end of each list. This extra item was not evaluated, and served as a means of minimizing eventual primacy and recency effects (Paivio, 1971). Therefore, each testing session consisted of the presentation of 14 slides. An equal number of food and animal items were accompanied by a picture or a phonological form. Types of presentation were diversified for each word length, and their order of presentation was randomized. Table 1 shows the details for the three lists.

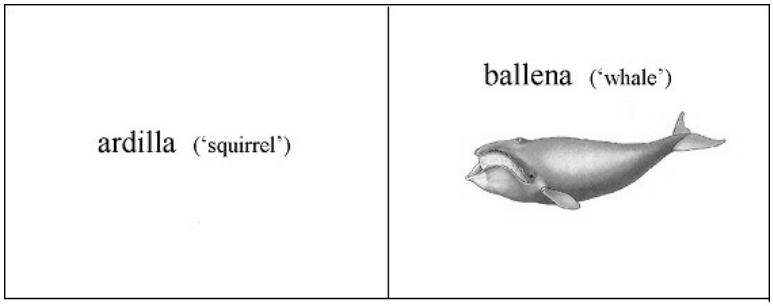

Figure 1: Slide samples for Conditions A (*ardilla*) and C (*ballena*).

Table 1: Details of the Item Lists

	List A			List B			List C		
Pn	Word	Cn	Lg	Word	Cn	Lg	Word	Cn	Lg
0	oso	C	–	cabra	C	–	vaca	C	–
1	alce	C	2;4	pulpo	A	2;5	tocino	B	3;6
2	pulga	B	2;5	apio	B	2;4	cisne	D	2;5
3	hongo	A	2;5	murciélago	A	4;10	miel	A	2;4
4	mariposa	D	4;8	morsa	D	2;5	rana	D	2;4
5	conejo	C	3;6	leche	B	2;5	durazno	C	3;7
6	maíz	A	2;4	mofeta	A	3;6	ardilla	A	3;7
7	zanahoria	C	4;9	fresa	C	2;5	queso	C	2;5
8	mosca	A	2;5	toronja	B	3;7	mapache	A	3;7
9	gallina	B	3;7	arándano	C	4;8	mariquita	C	4;9
10	cebolla	B	3;7	pavo	D	2;4	foca	B	2;4
11	berenjena	D	4;9	rábano	D	3;6	lechuga	D	3;7
12	pepino	D	3;6	ballena	C	3;7	mantequilla	B	4;11
0	oso	C	–	cabra	C	–	vaca	C	–

Notes:
Pn = Position of item in list; Cn = Condition of presentation; Lg = Length of written word (number of syllables;number of letters)
The first/last item is not measured or numbered since it is not considered in the analyses.

Procedure

Participants were tested in groups in a computer laboratory, each facing a computer screen. Time of presentation for each slide was pre-set at six seconds, with two-second blank-screen intervals. Presentation time had to be longer than the three seconds judged necessary for image generation to occur for concrete words (Paivio, 1966), and in order to be closer to classroom learning conditions.

Total presentation time of a list was one minute and 50 seconds. When sound was included, the pronunciation of the word started at the end of the first second.

Participants were informed that the purpose of the experiment was to verify how many items they could recall and that it was normal not to recall all of them. Recall was checked five minutes after the end of the presentation to avoid participants looking up words in their books or talking to one another in the meantime. Recall needed to be delayed by more than 30 seconds to reduce the recency effect (Paivio, 1966).

For the recall task, participants were given a sheet of paper with a list of the 12 items in English for which they had to write the Spanish equivalents. This procedure is intended for mixing recognition (all English words were on every slide) with a traditional way of assessing learning in the classroom, which consists of producing the L2 equivalent that was introduced in association with the word. They also had to circle the words they already knew or had already seen prior to the experiment. Ten minutes were allowed for the recall task.

Data analyses

One point was allowed for a correct answer, zero for an incorrect one. To be considered correct, a word had to be spelled correctly, or with mistakes that would not modify its pronunciation. For example, the Spanish word *fresa* 'strawberry' with a double *s* would be accepted, but not *fosa*, which is in fact closer to another item (*foca* 'seal'). Items for which the participants admitted previous knowledge were not considered in the analyses. An ANCOVA was run to determine the simultaneous influence of our independent variables as covariates (list used, word length measures, position and type of presentation). Pearson correlation coefficients were calculated to estimate the influence of word length measures and position of items on the scores. An alpha level of .05 was used for all statistical tests.

Precautions were taken to avoid the influence of variables that could affect results. Before addressing the research questions, it is important to comment on these precautions.

First, a two-sample *t*-test showed no significant difference in average recall between animal and food items ($t = 0.60$, $p = 0.55$). Second, because testing sessions took place over three different class periods, the number of participants for each session depended on attendance at these classes. Therefore, 45 of the participants did all three lists and five participants did two. Given the low percentage of participants who did only two lists and the fact that all three lists yielded statistically equivalent scores ($F = 0.20$; $p = 0.82$), it was not deemed that this would influence the results significantly. Finally, previous studies have shown better recall for items presented either at the beginning or at the end of a series (the aforementioned primacy and recency effects) than for items presented in the middle of the series. Having taken steps (extra item + longer

Table 2: Descriptive Statistics for the Presentation of Items

Type of presentation	Mean	Standard Deviation
Condition A: Word alone	27.93	4.87
Condition B: Word + Sound	5.22	4.31
Condition C: Word + Picture	22.85	4.05
Condition D: Word + Picture + Sound	18.29	4.24
Sound		
Without (= Conditions A + C)	25.39	3.13
With (= Conditions B + D)	11.75	2.99
Picture		
Without (= Conditions A + B)	16.57	3.28
With (= Conditions C + D)	20.57	2.95

Note: The values represent percentages.

recall delay) to minimize these positional effects seems to have worked, since Pearson product-moment correlations showed no significant effect of position on recall (r = -0.07, p = 0.71). The results are presented below in relation to the research questions identified earlier.

Results

The first two research questions concern the possible effects of extra stimuli on the recall of written words. Answering these questions requires comparing the recall averages for each type of presentation. The results of this comparison are presented in Table 2.

For external stimuli to show additivity, recall for the word alone (Condition A) should be significantly lower than for Conditions B, C and D. A first look at the data in the upper section of Table 2 suggests that this is not the case, since the average score for Condition A is the highest of all. Pairwise comparisons using Tukey simultaneous tests show that the average for Conditions C and D are statistically equivalent to that for Condition A (T = -0.79, p = 0.86 for C; $T = -1.42$, p = 0.50 for D). The average for Condition B, it should be noted, is significantly lower ($T = -3.52$, $p < 0.01$). This unpredicted aspect of our results is discussed below. Now, we will examine additivity for each of the two types of external stimuli.

Additivity of pictures

Our first research question asks whether recall is enhanced when pictures are added to mental images. In other words, for high-imagery words such as those used in this experiment, does the presence of a picture increase recall even though participants theoretically produce an internal image without the

picture? In order to verify this possibility, the recall average for words without a picture was compared to that for words accompanied by a picture. Combining the scores for Conditions A and B, where a picture was not present, yielded scores (M = 16.57, SD = 3.28) that are not significantly different (F = 0.75, p = 0.39) from the scores for the combined Conditions C and D (M = 20.57, SD = 2.95), where a picture was also shown to the participants. Therefore, these results tend to demonstrate that mental imagery is such a powerful aid for memory that when it has the time to occur, as was the case here, adding an actual picture did not significantly improve recall by the participants.

Additivity of phonology

Our second research question was whether recall is enhanced when external verbal stimuli are added to inner speech. Investigating this possibility requires combining the conditions in which the pronunciation of the item was provided to the participants (B + D) and comparing these scores to those for conditions in which inner speech was theoretically involved without auditory stimuli (A + C). This time, the presence of the spoken form influenced the scores significantly (F = 9.40, $p < 0.01$) but negatively (M = 11.75, SD = 2.99 with sound; M = 25.39, SD = 3.13 without sound). Therefore, the answer to our second question is negative. The absence of significantly better recall with the addition of the aural stimulus suggests that providing the pronunciation of a word does not add anything to inner speech in enhancing recall. On the contrary, adding the pronunciation results in a significantly lower recall. As mentioned before, this phenomenon will be treated in the discussion.

Influence of word length

The design of our experiment also permitted verification of whether word length influences recall. Our Pearson correlations demonstrate that word length is a significant predictor of recall for newly encountered words, yielding a correlation of -0.47 ($p < .01$) between recall and the number of letters and of -0.51 ($p < .01$) between recall and the number of syllables. An interesting observation is that for each supplementary syllable, recall is divided in half (M = 27.69% for 2 syllables; M = 14.88% for 3 syllables; M = 7.71% for 4 syllables). This observation can also be made from the results obtained recently by Campaña Rubio and Ecke (2001), who obtained 55.07% recall for two-syllable words, 27.54% for three-syllable words and 17.39% for four-syllable words.

Discussion and Conclusion

The results obtained in the present experiment suggest that, in a typical teaching situation where inner speech and imagery intervene, adding other types of

stimuli to the written word does not improve recall. Previous results that have suggested the contrary might have been due to a fast pace of presentation (most of the time well below one second) that is not normally used in the classroom and does not allow for mental imagery to take place. When students are given enough time to construct mental images, recall seems to depend more on other factors than on the presentation of a picture or illustration. This does not rule out the possibility that pictures are remembered better than words, because remembering pictures does not necessarily mean that the participant will associate them to the right Spanish words. Indeed, there were several instances in the present experiment of recall associated to wrong items. The same efficiency as for imagery can be attributed to inner speech, since adding the spoken form of the word does not increase recall.

As was mentioned earlier, an explanation is needed for the lower recall for conditions where the phonological form was provided. Words in Conditions B and D were not different from words in other conditions as far as length was concerned. One explanation is possible interference created by discrepancies between the form that is heard and the one created by inner speech. Such a contradiction between the actual stimulus and the self-generated one would more logically occur for phonology: mental images are retrieved based on already seen items, whereas inner speech would be generated without the item having been heard before. This phonological interference hypothesis could possibly explain results such as Brooks' (1967) that show better recall for complex messages in the auditory mode than in auditory mode accompanied by written text. Brooks concluded that this pattern of interference reflects the fact that verbal coding requires some of the same resources as auditory perception. More research is needed to shed light on the adverse effects in silent reading caused by discrepancies between the auditory stimulus and the self-produced phonological form.

Unfortunately, carrying out a study that focuses on learning makes it virtually impossible to compare different sets of stimuli for the same items and for the same individuals, since after the first set an item would no longer be a new item. Another limitation that comes with allowing for mental imagery and inner speech is that a longer time of presentation adds variability between the participants as to the number of possible mental retrievals of the items. The consequences of such rehearsals need to be explored.

The additivity hypothesis is based on and supported by many studies on word identification. The present experiment investigated addivity in a classroom context, where the participants produce inner speech and create mental images. Our results support earlier studies in underscoring the efficiency of self-created mental images for enhancing recall (McCabe, 1988; Gambrell and Jawitz, 1993; Bekerian and Dennett, 1997, for example), such that adding a picture to the word being taught might not add significantly to recall. The negative effect obtained

by the addition of the spoken form does not contradict Paivio's hypothesis, since a phonological form was indeed already present for Conditions A and C, through inner speech. As was stated earlier, conflicting information between the two phonological forms (the self-generated and the one being heard) is a possible explanation. This suggests that inner speech may not be as powerful as self-generated images for enhancing recall.

Our results further support the claim that the most obvious variable influencing recall for new vocabulary is word length, along with other individual variables that further research could help identify. For example, in this study, discussions with participants after the testing sessions suggested that recall is, as mentioned earlier, enhanced by linking the newly encountered word to already existing knowledge. For instance, some participants admitted remembering words because they looked like their own name or that of a relative (for example, *Renée*/*rana*) or a place with which they created a mental association. Looking at the items that were the most recalled (*fresa* 'strawberry' = 58.5%, *hongo* 'mushroom' = 42.9%, and *queso* 'cheese' = 42.2%), it is possible that, for instance, *fresa* 'strawberry' might have been associated with the beverage *Fresca* or with the English word *fresh*, easily associated with that fruit. Also, while as mentioned an attempt has been made to eliminate cognates, the Spanish word *hongo* 'mushroom' could nevertheless have been functioning as a cognate if compared to *fungus* by especially observant participants, as with the pair *queso* 'cheese'. Another example is that of a participant who also mentioned remembering *berenjena* because its English equivalent, *eggplant*, used to be a racial slur in the United States. This tends to further support the accepted notion that individual knowledge and learning strategies play an important role in recall, and the most recalled items seem to be those which allow for easy associations with already possessed knowledge.

This study raises interesting questions about the nature of applying scientific research to education. Research in psychology that deals with cognitive processes is often conducted under conditions that do not reflect those in the classroom. In the quest for more efficient pedagogical methods, educators have to be careful when considering evidence coming from research studies. This study suggests that for teaching isolated words, a good strategy consists of encouraging students to create mental associations for remembering new words, and to practice mental rehearsal. Mental imagery and inner speech being useful tools for recall, the emphasis in teaching should be put on longer words and abstract words more than on diversifying the stimuli used in the presentation.

Notes

The author wishes to thank Dr. Rich Einsporn for his help with the statistical analyses.

[1] Another recent explanation linking the picture superiority effect to the reduction of the cognitive load was formulated by Sweller and his colleagues (Mousavi, Low and Sweller, 1997, for example).

[2] Among the rare exceptions to this trend, Bousfield, Esterson and Whitmarsh (1957) showed better recall for pictures and words than for words alone, and for colour pictures than for black and white ones, arguing for the use of "an increased number of simultaneously presented additional signs of the objects denoted by these words" (p. 167).

References

Anderson, J.A. 1978. "Arguments concerning representations for mental imagery." *Psychological Review*, 85, pp. 249–277.

Avons, S.E., K.L. Wright and K. Pammer. 1994. "The word-length effect in probed serial recall." *Quarterly Journal of Experimental Psychology*, 47, pp. 207–231.

Baddeley, A.D. 1986. *Working Memory*. Oxford: Oxford University Press.

Baddeley, A.D., N. Thompson and M. Buchanan. 1975. "Word length and the structure of short-term memory." *Journal of Verbal Learning and Verbal Behavior*, 14, pp. 575–589.

Barsalou, L.W. 1999. "Perceptual symbol systems." *Behavioral and Brain Sciences*, 22, pp. 577–660.

Bekerian, D.A. and J.L Dennett. 1997. "Imagery effects in spoken and written recall." In *Intersections in Basic and Applied Memory Research*, D. Payne and F. Conrad (eds.). Mahwah, NJ: Lawrence Erlbaum, pp. 279–290.

Bousfield, W.A., J. Esterson and G.A. Whitmarsh. 1957. "The effects of concomitant colored and uncolored pictorial representations on the learning of stimulus words." *Journal of Applied Psychology*, 41, pp. 165–168.

Brooks, L. 1967. "The suppression of visualization by reading." *Quarterly Journal of Experimental Psychology*, 19, pp. 289–299.

Campaña Rubio, E.B. and P. Ecke. 2001. "Un estudio experimental sobre la adquisición y recuperación (parcial) de palabras en una lengua extranjera." In *Memorias del V Encuentro Internacional de Lingüística en el Noroeste*, G. López Cruz and M. Morúa Leyva (eds.). Hermosillo, Mexico: Editorial Unisón, pp. 63–84.

Cowan, N., L. Day, J.S. Saults, T.A. Keller, T. Johnson and L. Flores. 1992. "The role of verbal output time in the effects of word length on immediate memory." *Journal of Memory and Language*, 31, pp. 1–17.

Dyer, F.N. 1973. "The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes." *Memory and Cognition*, 1, pp. 106–120.

Gambrell, L.B. and P.B. Jawitz. 1993. "Mental imagery, text illustrations, and children's story comprehension and recall." *Reading Research Quarterly*, 28, pp. 264–276.

Gathercole, S.E. and A.D. Baddeley. 1990. "The role of phonological memory in vocabulary acquisition: A study of young children learning new names." *British Journal of Psychology*, 81, pp. 439–454.

Gildea, P.M., G.A. Miller and C.L. Wurtenberg. 1990. "Contextual enrichment by videodisc." In *Cognition, Education, Multimedia: Exploring Ideas in High Technology*, D. Nix and R. Spiro (eds.). Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1–30.

Kaufmann, G. 1996. "The many faces of mental imagery." In *Stretching the Imagination: Representation and Transformation in Mental Imagery*, C. Cornoldi, R.H. Logie, M.A. Brandimonte, G. Kaufmann and D. Reisberg (eds.). New York: Oxford University Press, pp. 77–118.

Kosslyn, S.M., W.L. Thompson and N.M. Alpert. 1997. "Neural systems shared by visual imagery and visual perception: A positron emission tomography study." *NeuroImage*, 6, pp. 320–334.

Lee, Y.-A., J.-O. Kim, K. Binder, A. Pollatsek and K. Rayner. 1999. "Activation of phonological codes during eye fixations in reading." *Journal of Experimental Psychology: Human Perception and Performance*, 25, pp. 948–964.

Lupker, S.J. 1985. "Relatedness effects in word picture and naming: Parallels, differences, and structural implications." In *Progress in the Psychology of Language*, Vol. 1, A.W. Ellis (ed.). London: Lawrence Erlbaum Associates, pp. 109–142.

Madigan, S. 1983. "Picture memory." In *Imagery, Memory and Cognition*, J.C. Yuille (ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 65–89.

Marschank, M. and C. Cornoldi. 1991. "Imagery and verbal memory." In *Imagery and Cognition*, C. Cornoldi and M.A. McDaniel (eds.). New York: Springer-Verlag, pp. 133–182.

Matz, K-D, J. Teschmer and G. Weise. 1988. "Angewandte Fremdsprachenpsychologie und ihr Beitrag für die Effektivierung des Lernens und Lehrens von Fremdsprachen." *Deutsch als Fremdsprache*, 4, pp. 224–229.

McCabe, A. 1988. "Effect of different contexts on memory for metaphor." *Metaphor and Symbolic Activity*, 3, pp. 105–132.

McGuigan, F.J. 1970. "Covert oral behavior during the silent performance of language tasks." *Psychological Bulletin*, 74, pp. 309–326.

Mellet, E., L. Petit, B. Mazoyer, M. Denis and N. Tzourio. 1998. "Reopening the mental imagery debate: Lessons from functional anatomy." *NeuroImage*, 8, pp. 129–139.

Mintzer, M.Z. and J.G. Snodgrass. 1999. "The picture superiority effect: Support for the distinctiveness model." *American Journal of Psychology*, 112, pp. 113–146.

Mousavi, S.Y., R. Low and J. Sweller. 1997. "Reducing cognitive load by mixing auditory and visual presentation modes." *Journal of Educational Psychology*, 87, pp. 319–334.

Paivio, A. 1965. "Abstractness, imagery, and meaningfulness in paired-associate learning." *Journal of Verbal Learning and Verbal Behavior*, 4, pp. 32–38.

Paivio, A. 1966. "Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality." *Canadian Journal of Psychology*, 20, pp. 378–387.

Paivio, A. 1971. *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.

Paivio, A. 1975. "Coding distinctions and repetition effects in memory." In *Psychology of Learning and Motivation*, Vol. 9. G. Bower (ed.). New York: Academic Press, pp. 179–214.

Paivio, A. and K. Csapo. 1973. "Picture superiority in free recall: Imagery or dual coding?" *Cognitive Psychology*, 5, pp. 176–206.

Paivio, A., M. Walsh and T. Bons. 1994. "Concreteness effects on memory: When and why?" *Journal of Experimental Psychology: Learning Memory and Cognition*, 20, pp. 1196–1204.

Pylyshyn, Z.W. 1973. "What the mind's eye tells the mind's brain: A critique of mental imagery." *Psychological Bulletin*, 80, pp. 1–24.

Richardson, J.T.E. 1980. *Mental Imagery and Human Memory*. London: Macmillan.

Rosinski, R.R., R.M. Golinkoff and K. Kukish. 1975. "Automatic semantic processing in a picture-word interference task."*Child Development*, 46, pp. 247–253.

Slowiacek, M.L. and C. Clifton. 1980. "Inner speech and reading for meaning." *Journal of Verbal Learning and Verbal Behavior*, 19, pp. 573–582.

Stroop, J.R. 1935. "Studies of interference in serial verbal reactions." *Journal of Experimental Psychology*, 18, pp. 643–662.

Tindall-Ford, S., P. Chandler and J. Sweller. 1997. "When two sensory modes are better than one." *Journal of Experimental Psychology: Applied*, 3, pp. 257–287.

Trojano, L. and D. Grossi. 1994. "A critical review of mental imagery defects." *Brain and Cognition*, 24, pp. 213–243.

Underwood, G. and V. Batt. 1996. *Reading and Understanding*. Cambridge, MA: Blackwell.

Zhang, S., C.A. Perfetti and H. Yong. 1999. "Whole word, frequency-general phonology in semantic processing of Chinese characters." *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, pp. 858–875.