

Input Modality and Remembering Name-Referent Associations in Vocabulary Learning

Hossein Nassaji
University of Victoria

This experimental laboratory-based study investigated the role of input modality in remembering name-referent associations in learning nonsense words. Three groups of L2 learners of English attempted to learn and remember name-referent associations in three different conditions: auditory only (n = 26), visual only (n = 28) and dual-modality (auditory/visual) (n = 25). Immediate recall and recognition of name-referent associations revealed no significant differences between the visual and auditory conditions. However, both recall and recognition of paired associations improved significantly when dual modality was used. These results do not seem to support the superiority of one mode of presentation over another for remembering name-referent associations in vocabulary acquisition. However, they do seem to provide support for the dual-modality hypothesis in vocabulary learning and the idea that semantic representations of words benefit from referential connections in which both auditory and visually-based processing is involved.

Cette étude expérimentale en laboratoire s'intéresse au rôle que joue la modalité d'entrée pour aider à se souvenir d'associations nominales référentielles dans l'acquisition du vocabulaire. Trois groupes d'apprenants d'anglais langue seconde ont tenté d'apprendre des associations nominales référentielles et de s'en souvenir dans les trois contextes suivants : modalité auditive (n = 26), modalité visuelle (n = 28) et modalité double (visuelle et auditive) (n = 25). Le rappel immédiat et la reconnaissance des associations nominales référentielles n'ont pas révélé de différence significative entre le visuel et l'auditif. Cependant, le rappel et la reconnaissance de paires d'associations étaient considérablement améliorés lorsqu'une modalité double était utilisée. Ces résultats ne semblent pas soutenir la supériorité d'un mode de présentation par rapport à un autre pour que les sujets se souviennent des associations nominales référentielles dans l'acquisition du vocabulaire. Cependant, ils semblent encourager l'hypothèse selon laquelle la double modalité est importante pour l'acquisition du vocabulaire, et l'idée que les représentations sémantiques des mots peuvent bénéficier de connections référentielles utilisant un processus basé sur le visuel et l'auditif.

Address for correspondence: Department of Linguistics, University of Victoria, P.O. Box 3045, Victoria, BC, V8W 3P4. E-mail: nassaji@uvic.ca.

Introduction

Current theories of memory suggest that there are separate mechanisms underlying auditory and visual information processing. One current view, the separate-streams hypothesis (Penney, 1980, 1989), suggests that visually and aurally presented verbal materials are processed and retained independently. According to Penney (1989, p. 399) “the processing of auditory and visually presented verbal items is carried out separately in short-term memory (the separate-streams hypothesis).” In this view, auditory and visual modes of presentation possess different characteristics, which differentially affect both encoding and retrieval processes. Penney suggests that the way learners organize recall is affected by the modality of the recall task.

The idea that information is processed in memory along different modalities is also central to Baddeley’s (1994) model of the working memory system of humans. Baddeley proposes that memory is composed of three components: a phonological loop component, a central executive component and a visuo-spatial sketchpad component. The phonological loop component processes speech and auditory-based information, and is also responsible for the phonological representations of visually represented materials. The executive component serves as a control mechanism, regulating and coordinating the overall processing and storing of information. The visuo-spatial sketchpad component represents and processes visual/imagery and spatial information. Based on this model, working memory has specialized systems for the perception, representation and retention of visually and auditorily presented information.

A wide range of experimental research has been carried out in the field of cognitive psychology on the effects of modality of presentation on human memory (Beaman, 2002; Beaman and Morton, 2000; Bird and Williams, 2002; Crowder, 1986; Dean, Yekovich and Gray, 1988; Frankish, 1985, 1995; Glensberg and Swanson, 1986; Greene, 1985). These studies have shown that the manner in which information is received affects its representation and learning in both short-term and long-term retrieval and retention. A number of studies have shown that auditory mode of presentation results in better immediate recall of verbal materials (Avons and Philips, 1980; Engle, Mabley and Linda, 1976; Penney, 1974, 1975, 1980). This effect has been found particularly with reference to learning verbal materials in serial recall tasks, and has been explained in terms of the robustness of the auditory sensory store in maintaining a phonological trace of auditorily presented words (Penney, 1989). Some studies suggest that while auditory presentation may lead to better short-term retention, visual presentation may be more effective for long-term retention and retrieval (Dean, Yekovich and Gray, 1988; Engle *et al.*, 1976; Krisner, 1974). Based on recognition data of visually and aurally presented words, Dean *et al.* (1988) found that a significantly greater number of words were remembered at

long recognition intervals when the words were presented visually rather than auditorily. The researchers concluded that there seems to be a link between the visual component of stimuli and their long-term retention.

L2 Studies on input modality

A number of studies have investigated the effects of input modality in L2 learning (Johnson, 1992; Johnson and Newport, 1989; Kelly, 1992; Lee, 1998; Lund, 1991; Murphy, 1997; Shih and Alessi, 1996). Johnson (1992) and Murphy (1997), for example, examined the effect of modality on learner performance on grammaticality judgment tasks. Johnson replicated Johnson and Newport's (1989) study on the critical period hypothesis, in which aural grammaticality judgment tasks had been used. Johnson examined whether a change in the modality of presentation — written instead of oral tasks — made a difference in the processing of these tasks. She found that written grammaticality judgment tasks elicited better performance than the aural tasks used in Johnson and Newport (1989). Murphy (1997) conducted a similar study with both L1 and L2 learners, and confirmed Johnson's results by finding that learners' responses were both slower and less accurate when the grammaticality judgment task was presented aurally as opposed to visually. Lund (1991) found that when L2 learners of German were presented with a text visually they recalled more details of the text than when they were presented with the text auditorily.

Wynne (2001) investigated the effects of modality on processing the form-meaning relationship in input. Motivated by VanPatten (1990), who had found, based on aural tasks, that Spanish L2 learners had difficulty attending to form and meaning at the same time, Wynne examined whether results would change if written tasks were used instead of aural tasks. Using both aural and written tasks, Wynne found that when the input was aural, there was a significant difference between a task in which the learners had to pay attention to content only and a task in which they had to attend to both content and form. When the input was written, no significant differences were found between the two tasks. She concluded that modality of presentation is a variable that has important effects on learner performance when processing linguistic input and that processing aural input is more difficult and needs considerably more attentional resources than processing written input.

A few studies have also examined the effects of single modality (either visual or auditory) versus dual modality (both visual and auditory). Kelly (1992) examined whether memory for foreign vocabulary was enhanced when learners read a text that contained the words versus when they both read and listened to the text. In a pilot study, Kelly had found a slightly stronger, but not significant, effect for the reading-only group on an immediate visual test, but a higher retention performance for the dual-modality group on delayed visual and aural tests. In a follow-up study, the researcher found a better and

significant immediate effect for the reading group on an immediate visual test but no better performance on a delayed visual test. On the delayed aural test, the performance was better in the dual-modality condition than in the reading-only condition. Kelly concluded that it seems that “the ear does assist the eye in the long-term retention of lexis” (p. 142). The results also suggest that the memory of foreign words depends not only on modality of presentation but on the type of memory measures used. Baltova (1999) also found that exposing L2 learners to words through bimodal videos combining sound, text and visual information enhanced both their understanding and learning of vocabulary from authentic texts.

While some research suggests that simultaneous presentation of information through different modalities may be helpful (see also Bird and Williams, 2002; Leahy, Chandler and Sweller, 2003; Mayer and Anderson, 1992), other studies suggest that adding to the modality of presentation may not assist learning (Pichette, 2002; Singer, 1980; Solman, Singh and Kehoe, 1992; Wu and Solman, 1993). Some studies with L1 children have demonstrated that the technique of presenting new words along with pictures does not help them in learning new words (Singh and Solman, 1990; Solman *et al.*, 1992; Wu and Solman, 1993, for example). It has been suggested that associating a word with pictures would divert the learner’s attention from the printed word, hence leading to poor processing of the word for learning (Wu and Solman, 1993). In a recent study with L2 learners, Pichette (2002) examined whether including pictures and pronunciation improved L2 learner recall performance of concrete words in Spanish. Testing the learning of the words in four modality conditions (word alone, word plus sound, word plus picture, and word plus sound and picture), Pichette found that not only was memory not improved by adding to the modality of presentation, but at times additional modality had negative effects, apparently interfering with learner performance. He found a higher recall performance for the word-only condition than the three other conditions, and a lower recall performance for the word/picture/sound condition than for the word/picture condition.

Thus, although it may be assumed that the presentation of information through more than one modality improves its learning, empirical evidence in this area is inconclusive, suggesting that further research is warranted into the effects of dual versus single modality in L2 learning. Moreover, it is important to note that although there are a few studies in SLA on the role of visual versus aural modes of presentation, these studies have mostly focused on the speed of processing or comprehension of linguistic information such as syntactic and semantic information rather than on learning. Current theoretical accounts of memory for language learning suggests that the degree with which new words are phonologically or visually represented in memory affects, and is linked to, the learning and retention of those words (Ellis, 2001). However, few

empirical studies have investigated these effects in L2 word learning. Learning new words in a language also requires that the learner learn and remember not only the word but the relationship established between the word and its referent (Duyck, Szmalec, Kemps and Vandierendonck, 2003), so that when the word is encountered, it elicits the referent and when the referent is encountered it elicits the word (Paivio, 1971; Steinberg, 1982; Stern, 1983). Thus, the present study investigated the effects of modes of presentation on remembering such word-referent associations with nonsense words. Three main research questions were formulated to examine these effects:

1. Does the kind of input modality, visual or auditory, have an effect on remembering name-referent associations of nonsense words?
2. Do subjects display any differential performance when different kinds of memory measures are used (recognition or recall)?
3. Are there any differential effects when single modality or dual modality is used?

Method

Subjects

Subjects were 79 adult university students who voluntarily participated in the study. All were students who were enrolled in general EFL (English as a Foreign Language) classes in a university context. Learners shared the same first language background, all being native speakers of Farsi. They were all males, ranging in age from 19 to 24 (Mean = 19.97). Based on their class placements, the subjects were assumed to be high-intermediate EFL learners. They reported no history of auditory or visual problems at the time of the experiment.

Materials

The experimental words designed for the study consisted of 24 unknown labels which the subjects were to learn as names for specific known objects. The use of real words was avoided in the study in order to minimize the confounding effects of learners' previous vocabulary knowledge and the frequency of real words in input (Kirsner, 1994). Therefore, instead of real words, a novel-word paradigm was used. Novel words have been widely used in psycholinguistic research to investigate lexical processes or to measure the effects of various input modalities on new word learning (Bird, Gaskell, Babineau and Macdonald, 2000; Bird and Williams, 2002; Monsell, 1985; Senechal, 1997). The assumption has been that under certain experimental conditions the learning of such words may provide evidence for the learning of new word forms (Bird and Williams, 2002, Monsell, 1985).

The labels used in the study denoted names of objects rather than actions. They were constructed to be monosyllabic, having consonant-vowel-consonant (CVC) structures (e.g., MEJ, NAZ, TEM) so that the words could be presented and processed quickly at one trial. The phonemes comprising the words were those shared between Farsi, the subjects' L1, and English, their L2. Farsi is an alphabetic language with a writing system different from that of English in terms of both physical shape and the way the letters combine to form an orthographic pattern. However, Farsi native speakers make phonemic distinctions similar to those made by native speakers of English, except for the fact that Farsi does not have the phonemes /ð/ and /θ/. The use of these two phonemes was therefore avoided in the test items.

Three experimental blocks were constructed, each consisting of eight labels randomly paired with eight specific referents, which were coloured drawings of common objects. The number of pairs in each block was determined based on the results of a pilot study with a group of subjects with almost the same educational and age characteristics. It was found that subjects' memories began to be seriously challenged as soon as they were exposed to more than seven or eight associative pairs. The experimental blocks of eight associative pairs of name-referents were presented to the subjects under three different modality conditions: auditory, visual, and auditory/visual.

Procedure

Subjects in each condition were first told about the general purpose of the study and then completed a short background questionnaire including questions about their age, gender and any visual or hearing problems. They were then presented the experimental blocks of eight name-referent associations in one of three different modality conditions: visual, auditory, and visual/auditory conditions. In the visual condition, the pictorial referents were presented to learners in association with written words. In the auditory condition, they were associated with oral words. In the dual-modality condition, they were associated with both oral and written words. The experiments took place in a language laboratory. The pairs of word-referents were displayed, with the use of a projector, in the centre of a large screen in front of the room (see Figure 1). Each presentation condition was preceded by three practice items. The rate of presentation in each condition was six seconds per pair, which has been considered sufficient time for subjects to view and register an association between the word and its referent (see Pichette, 2002). This rate was controlled by a computerized timer attached to the projector and was kept constant across conditions. Each of the subjects sat in an individual booth equipped with headphones and a tape recorder. Instructions for the experiment were provided both in the subjects' L2, English, and in their L1, Farsi. In the visual condition, the instructions were as follows:

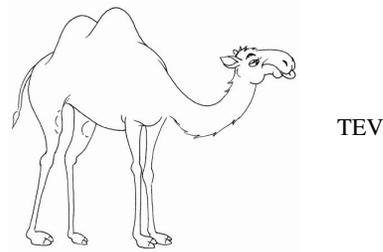


Figure 1: An Example of a Name-Referent Pair for the Visual Modality Condition

You will see a series of foreign words one by one, each associated with a corresponding picture appearing on the screen. The word you see represents the name of the picture. You have to attend to the word and the corresponding picture and try to learn the word as the name of that picture. You have to be able to remember the name of the picture when you are shown the picture and asked to name it.

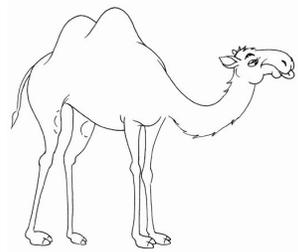
In the auditory condition, the instructions were as follows:

You will hear a series of foreign words one by one, each associated with a corresponding picture appearing on the screen. The word you hear represents the name of the picture. You have to attend to the word and the corresponding picture and try to learn the word as the name of that picture. You have to be able to remember the name of the picture when you are shown the picture and asked to name it.

In the dual-modality condition, the same instructions were given except that the subjects were told that they would simultaneously see and hear the words.

Testing procedures

Subjects' memory of the paired associations was tested immediately after each block presentation. Half of the items in each block were tested by recall and the other half by recognition tests. Because there were three blocks of eight name-referent associations, altogether twelve items were tested by recall (3 blocks \times 4 items) and twelve items by recognition tests (3 blocks \times 4 items). Order of recognition and recall was counterbalanced across modality conditions. That is, in each condition half of the subjects took the first four test items in the form of a recognition test while the other half took them in the form of a recall test. The modality of the testing condition was kept parallel to that of the learning condition. For the recognition tests, pictures in each experimental block appeared one by one and in random order on the screen, each associated with four names listed as a, b, c and d, one of which was the name they had to learn in the learning condition (see Figure 2 for an example). Subjects had



- a. LEB
- b. SOJ
- c. TEV
- d. KAG

Figure 2: An Example of a Test Item for the Visual Recognition Condition

to indicate which of the names referred to the picture next to it by circling the corresponding letter of the alphabet on their answer sheets. In the visual recognition task, they viewed both the picture and the words on the screen. In the aural recognition task, they viewed the pictures on the screen but heard the words in their headphones. In the auditory/visual recognition task, they viewed the pictures while they simultaneously heard and saw the word. In order to minimize the possibility of subjects' choosing the right answer by eliminating the ones that they had not encountered, the four options in each test item were chosen among the labels learners had already encountered in the modality conditions. The same procedure was repeated for the recall tests. Immediately after each block of eight paired stimuli, the pictures in each block appeared one by one and in random order on the screen. This time, however, instead of being given names and being asked to recognize the names of the pictures, subjects were asked to recall the names. In the visual-recall task subjects had to recall the name of the picture and write it on their answer sheets. In the auditory recall task they had to recall the name and say it into their microphone, and in the aural/visual recall task they had to either say the name into their microphones or write it on their answer sheets.

Design

The study used a 3 (Modality: auditory, visual, auditory/visual) \times 2 (Measure: recognition, recall) mixed factorial design. Each of the three groups of subjects participating in the study was randomly assigned to one of the three modality conditions. Thus, mode of presentation was a between-group factor. Subjects' memories were tested using both recognition and recall tests. Thus, memory measure was a within-group factor.

Data analysis

In analysing the data, one point was assigned for each correctly recognized or recalled name-referent association. The number of correct responses was tallied for the testing in each input modality condition. Because there were twelve

items tested by recall and twelve items by recognition tests in each condition, possible scores for each test ranged from 0.00 to 12.00. Because there were three modality conditions, this yielded six sets of scores: three for recognition (auditory recognition, visual recognition and auditory/visual recognition), and three for recall (auditory recall, visual recall, and auditory/visual recall). Descriptive statistics were used to calculate the mean and standard deviation for each of six condition conditions. A multivariate repeated measure of variance (MANOVA) including tests for simple main effects and interaction effects and a post hoc Scheffé test of multiple comparisons were then applied. The repeated measure included one between-subject factor, which was the modality condition with three levels: auditory, visual, and dual-modality condition; and one within-subject factor, which was the kind of measure used, with two levels: recognition and recall.

Results

Table 1 displays the means and the standard deviations of the recognition and recall tests for the three different modality conditions. The mean score for recognition tests in the auditory/visual condition was 7.04, which was significantly higher than the mean score for either visual recognition (5.25) or auditory recognition (5.42). Similarly, the mean score for the recall tests in the auditory/visual condition was 4.88, which was significantly higher than the mean score for either visual recall (3.82) or auditory recall (4.03).

Table 1: Means and Standard Deviations of Immediate Recognition and Recall

Measures	Conditions	<i>M</i>	<i>SD</i>	<i>N</i>
Recognition (0–12)	Auditory	5.42	1.39	26
	Visual	5.25	1.35	28
	auditory/visual	7.04	1.36	25
	Total	5.87	1.57	79
Recall (0–12)	Auditory	4.03	1.14	26
	Visual	3.82	1.24	28
	auditory/visual	4.88	1.05	25
	Total	4.22	1.22	79

The results of the MANOVA given in Table 2 showed a significant main effect for the presentation condition ($F(2, 76) = 10.07, p < .0001$), suggesting that there was a significant difference among the different modality conditions.

However, the results of a post hoc Scheffé test, given in Table 3, revealed that this difference was not due to the difference between visual and auditory modalities. Although learners tended to perform slightly better in the auditory condition than in the visual condition, the pair-wise comparison of the different

Table 2: Test of Between Subject Effects (Main Effect for Modality of Presentation)

Source	Type III SS	df	MS	F	p
Intercept	4061	1	4061	1329	.000
Modality of Presentation	61	2	30	10.07	.000

Table 3: Multiple Comparisons (post hoc Scheffé test)

Measures	Conditions		Mean Difference
Recognition	Auditory	Visual	.17
		Auditory/Visual	-1.62*
	Visual	Auditory/Visual	-1.79*
Recall	Auditory	Visual	.21
		Auditory/Visual	-.85*
	Visual	Auditory/Visual	-1.06*

Note: * $p < .05$

conditions showed no significant differences between the visual and auditory modalities, but it did show a significant difference between the single (visual or auditory) and the dual modality (visual and auditory) condition.

These results suggests that learners were significantly more accurate in recognizing and recalling the name-referent associations when they received them both aurally and visually than when the received them either aurally or visually.

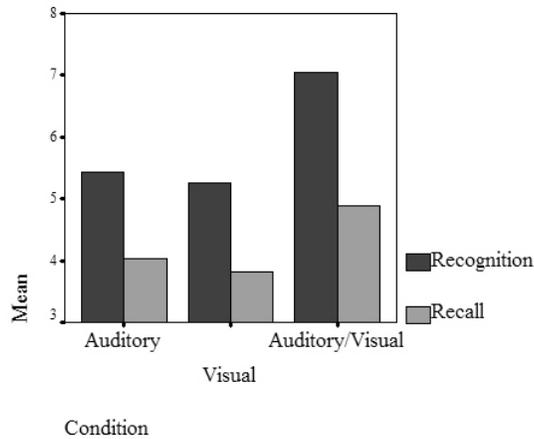
The results of a MANOVA also showed a significant main effect for types of measure: recognition versus recall. The total mean numbers of correct responses for the recognition and recall tests were 5.87 ($SD = 1.57$) and 4.22 ($SD = 1.22$), respectively: ($F(1, 77) = 675, p < .0001$). In all three conditions, subjects were significantly more accurate at recognizing than recalling the words after the presentation. There was also an interaction effect for the type of measure and presentation condition which was statistically significant ($F(2, 76) = 15.13, p < .0001$). These results are presented in Table 4 and are graphically displayed in Figure 3. As can be seen, the difference between the recognition and recall in the dual-modality condition was greater than the difference in recognition and recall in the other two modality conditions. Recognition was higher than recall in the dual-modality condition than in the other two conditions.

Discussion and Conclusion

The present experimental study examined whether different types of modality of presentation have any differential effects on remembering and retaining name-referent associations when subjects are learning nonsense words for common

Table 4: MANOVA of Main Effect for Type of Measure and Interaction between Type of Measure and Presentation Condition

Effect		Value	<i>F</i>	<i>df</i>	Error <i>df</i>	<i>p</i>
Type of Measure	Wilks' Lambda	.101	675	1	76	.000
Type of Measure × Condition	Wilks' Lambda	.715	15.13	2	76	.000

**Figure 3:** Effects of Modality of Presentation on Recognition and Recall

objects in a laboratory setting. The results showed that subjects' immediate recall and recognition of name-referent associations did not differ between visual and auditory modalities. Although subjects who received the names aurally tended to recall and recognize them slightly better than those who received them visually, a pair-wise comparison of the two conditions showed that this increased effect was not statistically significant. In general, these results do not seem to support the superiority of one mode of presentation over another for learning name-referent associations in vocabulary acquisition. The lack of significant effect of modality for single modes of presentation may also be explainable in terms of the interconnectedness of human memory systems and the idea that information encoded in one mode can be recoded into another in memory (Penney, 1993). It is possible that the subjects engaged in such a cross-modality recoding strategy and that the use of this strategy may have balanced off the possible effect of one modality of presentation over another.

The results of the single- versus dual-modality conditions showed that learners recalled and recognized the name-referent associations significantly

better when the words were presented both visually and aurally in comparison to when they were presented either visually or aurally. Results also showed a significant interaction between learners' recall and recognition and presentation conditions. Learners in the dual-modality condition showed a stronger ability to accurately recognize the names than to recall them. These findings provide empirical support for the dual-modality hypothesis in new word learning, and are consistent with several theoretical accounts in cognitive psychology that suggest that dual-modality presentation is more effective than single-modality presentation (e.g., Kirschner, 2002, Paas, Renkl and Sweller, 2003; Paivio, 1971, 1986, 1991; Sweller, 1993). The advantage of dual modality is consistent with the dual-coding theory of information (Paivio, 1971, 1986, 1991). According to the dual-coding theory, human cognition consists of two different, but interconnected, processing sub-systems, one visual (imagery system) specialized for processing and representing nonverbal information such as pictorial information, and the other a verbal system responsible for representing linguistic information. The two systems are activated independently, but the interconnections between the two systems allow for the dual coding of information. In this view, the concurrent and simultaneous use of both modalities leads to the construction of two mental representations and memory traces of the same information, hence enhancing recall and recognition of information (Paivio, 1991).

The facilitative effect of dual modality is also consistent with the cognitive load theory of learning (CLT) (Kirschner, 2002; Paas *et al.*, 2003; Sweller, 1993). The CLT theory is based on the assumption that learning involves working-memory capacity. It suggests that our cognition is based on a cognitive architecture that has a limited working-memory capacity and independent visual and aural processing units, and that the limitations of working memory entail that learners should make optimal use of this cognitive system by decreasing the extraneous mental load imposed on learning. The theory assumes that two characteristics would affect the learner's cognitive load and the learner's final performance and mental effort allocated to the task. One characteristic relates to the learner's cognitive ability, and the other relates to the nature of the task or the environment in which or the manner by which the task is presented. According to this theory, one way of decreasing this mental load is through multiple coding of information. This includes presenting materials in various supporting formats, providing examples for abstract cases, or making simultaneous use of different modalities for presentation of information. The use of both the visual and auditory mode of presentation has been shown to lead to the reduction of the learner's cognitive load or the circumvention of the limitations of the learner's working memory, and hence to learning more content (Tindall-Ford, Chandler and Sweller, 1997) and vocabulary from the text (Plass, Chun,

Mayer and Leutner, 2003) than when only one mode of presentation, either visual or auditory, has been used.

It is important to point out that while the superior recall and recognition performance found in this study for the dual-modality presentation is consistent with the findings of several studies on the effect of dual modality, it is not consistent with Pichette's (2002) results. Pichette found that subjects who received Spanish words with aural and visual cues did not recall the words better than those who received them without aural cues. As the researcher suggested, this lack of effect may have been due to the difference between the pronunciation generated by the learner and the actual pronunciation of the word. In the present study, there could have been less interference from the difference between the pronunciation generated through grapheme-phoneme correspondence rules and the actual pronunciation of the words. Moreover, research has shown that the effects of adding to the mode of presentation are stronger when learners are learning phonologically unfamiliar words than when they are learning phonologically familiar words (Bird and Williams, 2002). This effect could be due to the fact that when words are unfamiliar, learners may not be able to establish a phonological representation of the word based on its sounds only. Therefore, when the learners are learning completely new labels, as in this study, semantic and phonological representations of words may benefit from referential connections in which both auditory and visually based processing is involved.

Implications, Limitations and Suggestions for Further Research

It is important to note that this study was conducted under experimental conditions. Therefore, the implications of the results for classroom teaching must be treated with caution. Indeed, the purpose of the study was not to examine the effectiveness of different techniques of classroom teaching for vocabulary learning. It was to conduct an experimental study to determine whether or not there is any particular difference between modality conditions for remembering name-referent associations. This purpose, however, need not prevent one from taking the results as having relevance for language teaching. For example, the increased performance in the dual-modality condition may be taken as support for the idea that a bimodal vocabulary teaching method may be more advantageous than a single modality approach to learning and remembering new words (see Pouwels, 1992). One way of achieving this goal is through supporting listening and vocabulary learning tasks with textual and pictorial annotations using multimedia (Jones and Plass, 2002; Plass *et al.*, 2003), closed-captioned television programs in the classroom (Vanderplank, 1993), or adding written subtitles to videos while L2 learners are performing listening tasks (Baltova, 1999). The use of these techniques in language classrooms has been shown

to enhance both language comprehension and vocabulary learning (see also Goldman and Goldman, 1988; Vanderplank, 1992).

There are several other limitations that should be considered when interpreting the results. This study used monosyllabic nonsense words to name the pictures. This decision was made to control for the confounding effects of vocabulary knowledge and its frequency in input. However, the use of monosyllabic stimuli may be problematic in that not all words in a language are single-syllable and not all are so short or necessarily have CVC structure. Therefore, further studies using labels more representative of real words with more complex phonological and orthographic structures are needed. Future studies should also examine how modality of presentation may interact with differences in lexical characteristics such as length, frequency, and phonological and orthographic structures. These variables seem to be important variables mediating the effect of modality of presentation in vocabulary learning (see Pichette, 2002). Future studies should also examine the extent to which these effects are influenced by the individual learner's verbal ability (Plass *et al.*, 2003) or his/her perceptual learning styles and strategies (Ehrman and Oxford, 1990; Oxford, 1990; Reid, 1995).

The present study used a between-group research design, a design in which different groups of subjects were assigned to different modality conditions. Although the use of such a design allowed for the examination of the effects of modality when learners were to learn the same stimuli under different conditions, it does not control for the differences in the results arising from individual differences. Another possibility is to use a within-group research design, using the same groups of subjects in all conditions. However, care should be taken in the use of same-group designs because, although such a design may control for individual differences, it may be susceptible to test re-test or repetition effects if the same group of subjects is tested repeatedly under different conditions.

Note

I would like to thank Dr. Maggie Warbey for her comments on an earlier version of this article. I would also like to thank the three anonymous reviewers and the English-language editor of the *Journal* for their careful reading and very helpful comments.

References

- Avons, S.E. and W.A. Phillips. 1980. "Visualization and memorization as a function of display time and poststimulus processing time." *Journal of Experimental Psychology: Human Learning and Memory*, 6, pp. 407-420.
- Baddeley, A. 1994. "Working memory: The interface between memory and cognition." In D.L. Schachter and E. Tulving (eds.), *Memory systems*. Cambridge, MA: MIT Press, pp. 351-367.

- Baltova, I. 1999. "Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French." *Canadian Modern Language Review*, 56, pp. 31–48.
- Beaman, C.P. 2002. "Inverting the modality effect in serial recall." *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, 55A, pp. 371–389.
- Beaman, C.P. and J. Morton. 2000. "The effects of rime on auditory recency and the suffix effect." *The European Journal of Cognitive Psychology*, 12, pp. 223–242.
- Bird, E.K.-R., A. Gaskell, M.D. Babineau and S. Macdonald. 2000. "Novel word acquisition in children with down syndrome: Does modality make a difference?" *Journal of Communication Disorders*, 33, pp. 241–266.
- Bird, S.A. and J.N. Williams. 2002. "The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling." *Applied Psycholinguistics*, 23, pp. 509–533.
- Cowan, N. 1995. *Attention and memory: An integrated framework*. Oxford: Oxford University Press.
- Crowder, R.G. 1986. "Auditory and temporal factors in the modality effect." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, pp. 268–278.
- Dean, R.S., F.R. Yekovich and J.W. Gray. 1988. "The effect of modality on long-term recognition memory." *Contemporary Educational Psychology*, 13, pp. 102–115.
- Duyck, W., A. Szmalec, E. Kemps and A. Vandierendonck. 2003. "Verbal working memory is involved in associative word learning unless visual codes are available." *Journal of Memory and Language*, 48, pp. 527–541.
- Ehrman, M. and R. Oxford. 1990. "Adult language-learning styles and strategies in an intensive training setting." *Modern Language Journal*, 74, pp. 311–327.
- Ellis, N. 2001. "Memory for language." In P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 33–68.
- Engle, R.W., L.A. Mobley and A. Linda. 1976. "The modality effect: What happens in long-term memory?" *Journal of Verbal Learning and Verbal Behavior*, 15, pp. 519–527.
- Frankish, C. 1985. "Modality-specific grouping effects in short-term memory." *Journal of Memory and Learning*, 24, pp. 200–209.
- Frankish, C. 1995. "Intonation and grouping in immediate serial recall." *Applied Cognitive Psychology*, 9, pp. 5–22.
- Glenberg, A.M. 1984. "A retrieval account of the long-term modality effect." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, pp. 16–31.
- Glenberg, A.M. and N.G. Swanson. 1986. "A temporal distinctiveness theory of recency and modality effects." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, pp. 3–15.
- Goldman, M. and S. Goldman. 1988. "Reading with closed-captioned TV." *Journal of Reading*, 31, pp. 458–461.
- Greene, R.L. 1985. "Constraints on the long-term modality effect." *Journal of Memory and Learning*, 24, pp. 526–541.

- Gronlund, S.D. and R. Ratcliff. 1989. "The time course of item and associative information: Implications for global memory models." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, pp. 846–858.
- Johnson, J.S. 1992. "Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence." *Language Learning*, 42, pp. 217–248.
- Johnson, J.S. and E.L. Newport. 1989. "Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language." *Cognitive Psychology*, 21, pp. 60–99.
- Jones, L.C. and J.L. Plass. 2002. "Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations." *Modern Language Journal*, 86, pp. 546–561.
- Kelly, P. 1992. "Does the ear assist the eye in the long-term retention of lexis?" *IRAL*, XXX, pp. 137–145.
- Kirsner, K. 1994. "Implicit processes in second language learning." In N. Ellis (ed.), *Implicit and explicit learning of language*. San Diego, CA: Academic Press, pp. 283–312.
- Kirsner, K. 1974. "Modality differences in recognition memory for words and their attributes." *Journal of Experimental Psychology*, 102, pp. 579–584.
- Kirschner, P.A. 2002. "Cognitive load theory: Implications of cognitive load theory on the design of learning." *Learning and Instruction*, 12, pp. 1–10.
- Leahy, W., P. Chandler and J. Sweller. 2003. "When auditory presentations should and should not be a component of multimedia." *Applied Cognitive Psychology*, 17, pp. 401–418.
- Lee, J. 1998. "The relationship of verb morphology to second language reading comprehension and input processing." *Modern Language Journal*, 82, pp. 33–48.
- Monsell, S. 1985. "Repetition and the lexicon." In A. Ellis (ed.), *Progress in the psychology of language* (Vol. 2). London: Erlbaum, pp. 147–195.
- Murphy, V.A. 1997. "The effect of modality on a grammaticality judgement task." *Second Language Research*, 13, pp. 34–65.
- Oxford, R.L. 1990. *Language learning strategies: What every teacher should know*. Boston, MA: Heinle and Heinle.
- Paas, F., A. Renkl and J. Sweller. 2003. "Cognitive load theory and instructional design: Recent developments." *Educational Psychologist*, 38, pp.1–4.
- Paivio, A. 1971. *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Paivio, A. 1986. *Mental representation: A dual-coding approach*. New York: Oxford University Press.
- Paivio, A. 1991. "Dual coding theory: Retrospect and current status." *Canadian Journal of Psychology*, 45, pp. 255–258.
- Penney, C. 1974. "Order of report and interface effects in four channel bisensory memory." *Canadian Journal of Psychology*, 28, pp. 371–382.
- Penney, C. 1975. "Modality effects in short-term verbal memory." *Psychological Bulletin*, 82, pp. 68–84.
- Penney, C. 1980. "Order of report in bisensory verbal short-term memory." *Canadian Journal of Psychology*, 34, pp.190–195.

- Penney, C. 1989. "Modality effects and the structure of short-term verbal memory." *Memory and Cognition*, 17, pp. 398–422.
- Pichette, F. 2002. "Second-language vocabulary learning and the additivity hypothesis." *Canadian Journal of Applied Linguistics*, 5, pp.117–130.
- Plass, J.L., D.M. Chun, R.E. Mayer and D. Leutner. 2003. "Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities." *Computers in Human Behavior*, 19, pp. 221–243.
- Pouwels, J.B. 1992. "The effectiveness of vocabulary visual aids for auditory and visual foreign language students." *Foreign Language Annals*, 25, pp. 391–401.
- Reid, J.M. 1995. *Learning styles in the ESL/EFL classroom*. Boston, MA: Heinle and Heinle.
- Senechal, M. 1997. "The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary." *Journal of Child Language*, 24, pp.123–138.
- Singer, H. 1980. "Sight word learning with and without pictures: A critique of Arlin, Scott and Webster's research." *Reading Research Quarterly*, 15, pp. 290–298.
- Shih, Y. and S.M. Alessi. 1996. "Effects of text versus voice on learning in multimedia courseware." *Journal of Educational Multimedia and Hypermedia*, 5, pp. 203–218.
- Solman, R.T., N.N. Singh and E.J. Kehoe. 1992. "Pictures block the learning of sight-words." *Educational Psychology*, 12, pp. 143–153.
- Steinberg, D. 1982. *Psycholinguistics: Language, mind and world*. London: Longman.
- Stern, H.H. 1983. *Fundamental concepts of language teaching*. Oxford: Oxford University Press.
- Sweller, J. 1993. "Some cognitive processes and their consequences for the organization and presentation of information." *Australian Journal of Psychology*, 45, pp. 1–8.
- 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.
- Vanderplank, R. 1992. "Non-native speaker viewers' use of teletext subtitles in English: Some exploratory studies." In G. Davies and M. Hussey (eds.), *New technology in language learning*. Frankfurt am Main, Germany: Peter Lang., pp. 89–102.
- Vanderplank, R. 1993. "A very verbal medium: Language learning through closed captions." *TESOL Journal*, 3, pp. 10–14.
- VanPatten, B. 1990. "Attention to form and content in the input." *Studies in Second Language Acquisition*, 12, pp. 287–301.
- Wu, H.M. and R.T. Solman. 1993. "Effective use of pictures as extra stimulus prompts." *British Journal of Educational Psychology*, 63, pp. 144–160.
- Wynne, W. 2001. "Modality and attention to meaning and form in the input." *Studies in Second Language Acquisition*, 23, pp. 345–368.