Assessing the Effect of Lexical Variables in Backward Recall

Katherine Guérard and Jean Saint-Aubin
Université de Moncton

In a recent study, Bireta et al. (2010) suggested that when participants are required to recall lists of items in the reverse order, more attention is devoted to the recall of order at the expense of item information, leading to the abolition of item-based phenomena (the item and order trade-off hypothesis). In order to test the item and order trade-off hypothesis, we manipulated 4 lexical factors that are well known to influence item retention. The effects of word frequency, of lexicality, of semantic similarity, and of imageability were tested in forward and backward recall. All 4 phenomena were maintained in backward recall, which contradicts the item and order trade-off hypothesis. Instead, we suggest that backward recall might rely on semantic retrieval strategies.

Keywords: immediate memory, backward recall, long-term memory

The ability to retain sequential information is central to most everyday activities, as nicely illustrated by the notoriety of Lashley’s (1951) famous article on the problem of serial order in behavior (Bruce, 1994). Since Ebbinghaus (1885/1965), the study of order retention is one of the most researched topics in experimental psychology. The classic paradigm used in the study of order memory is the immediate serial recall paradigm, where participants are required to recall lists of items in their presentation order (e.g., see Baddeley, 1966; Murdock, 1968). Backward recall is a variant of the typical serial recall paradigm in which participants are asked to recall series of items in the reverse order. Such a task, which is used extensively as a measure of memory capacity, yields very different patterns of performance compared with forward serial recall, as evidenced by, among other means, the presence of separate scales for forward and backward digit span in Wechsler intelligence tests (see, e.g., Banken, 1985; Reynolds, 1997). However, although the extensive research using the serial recall paradigm has led to the development of a number of models (e.g., Baddeley & Hitch, 1974; Brown, Neath, & Chater, 2007; Nairne, 1990; Page & Norris, 1998), the processes that allow people to recall information in the reverse order are still difficult to explain by most models of memory (e.g., see Lewandowsky & Farrell, 2008).

One distinction that has been suggested in order to explain the different patterns in forward and backward recall performance is that between item and order retention (Bireta et al., 2010; see also DeLosh & McDaniel, 1996). Indeed, it is widely assumed that retaining information in serial recall requires memory for both the items’ identity and their order in the series (e.g., see Healy, 1974). Recently, Bireta et al. (2010; see also Surprenant et al., 2011) suggested that in backward recall, more attention is devoted to the recall of order information, and therefore less attention is given to the item dimension—what is known as the item and order trade-off hypothesis (see also DeLosh & McDaniel, 1996). Under this view, memory phenomena that mostly affect item retention should be abolished in backward recall.

A distinction between item and order retention is also invoked to account for the contribution of lexical variables to immediate serial recall performance. This contribution has been evidenced by, among other phenomena, the better recall of words over nonwords (see, e.g., Hulme, Maughan, & Brown, 1991) and of high- over low-frequency words (see, e.g., Roodenrys & Quinlan, 2000). As revealed by an analysis of item and order errors, these lexical variables influence item recall and have either no effect or a limited impact on order recall (see, e.g., Saint-Aubin & Poirier, 1999a). In the present study, we tested the item and order trade-off hypothesis by manipulating all major lexical variables known to affect item memory in backward and forward recall, that is, the effects of word frequency, lexicality, word concreteness, and semantic similarity, combined with an analysis of item and order errors.

**Backward Recall**

In backward recall, participants are asked to recall a series of items beginning with the last presented item back to the first presented item. Backward recall is usually contrasted with a condition of forward recall, where participants have to recall the items in the same order they were presented, beginning with the first presented item up to the last. The typical pattern of performance observed in backward recall is a pronounced advantage for the last items of the list, known as the recency effect, and a less marked advantage for the first items of the list, the primacy effect (see Li & Lewandowsky, 1995). This pattern of results contrasts with that observed in forward recall, which is characterized by a marked primacy effect and a more limited recency effect. A number of
Lexical Variables

Lexical variables refer to “word properties, such as word frequency and concreteness, which reflect differences in the way in which words are represented in long-term or lexical memory” (Stuart & Hulme, 2009, p. 157). Therefore, lexical variables include variables such as word frequency, lexicality, semantic similarity, and imageability—or concreteness. These variables are typically manipulated in order to investigate the role of long-term memory in immediate serial recall. The influence of these variables in immediate serial recall has traditionally been studied using forward serial recall instructions. More precisely, it is now well established that high-frequency words are better recalled than are low-frequency words (the word frequency effect; see, e.g., Hulme, Stuart, Brown, & Morin, 2003; Roodenrys & Quinlan, 2000; M. J. Watkins, LeCompte, & Kim, 2000); that words are better recalled than are nonwords (the lexicality effect; see, e.g., Hulme et al., 1991; Multhaup, Balota, & Cowan, 1996); that concrete, high-imageability words are better recalled than are abstract, low-imageability words (the effect of imageability or concreteness; see, e.g., Miller & Roodenrys, 2009; Romani, McAlpine, & Martin, 2008; Walker & Hulme, 1999); and that lists of semantically similar words are better recalled than are lists of dissimilar words (the effect of semantic similarity; see, e.g., Poirier & Saint-Aubin, 1995; Saint-Aubin, Ouellette, & Poirier, 2005; Tse, 2009). In addition, an analysis of errors showed that lexical variables affect recall mostly by modulating memory for item information with almost no effect on memory for order (see, e.g., Saint-Aubin & Poirier, 1999a). For example, Poirier and Saint-Aubin (1996) manipulated word frequency. They showed that high- and low-frequency words yielded a similar level of order errors but that the level of item errors was much higher for low- than for high-frequency words. They concluded that word frequency improved item memory, with no effect on the retention of order. This finding was also replicated with the effects of semantic similarity (e.g., Saint-Aubin & Poirier, 1999b) and of word imageability (e.g., Walker & Hulme, 1999). Among the four effects just discussed, only lexicality appears to affect order recall—but to a much smaller extent than item recall (see Saint-Aubin & Poirier, 2000).

According to the item and order trade-off hypothesis, item-based phenomena should be abolished in backward recall (Bireta et al., 2010). Consequently, under this view, lexical variables should have no effect in backward recall, because they mostly influence memory for item information. To the best of our knowledge, the influence of backward recall on the lexical effects has been investigated in only two studies, and their results are contradictory (Hulme et al., 1997; Walker & Hulme, 1999). Hulme et al. (1997) manipulated word frequency in backward and forward recall. In line with the item and order trade-off hypothesis, they found that the word frequency effect disappeared in backward recall. However, two closed pools of eight items were used, and it is likely that participants did not rely heavily on the item dimension in this task. It is therefore unclear whether the abolition of the word frequency effect in backward recall can be accounted for by a change in weight between the item and order dimensions. In another study, Walker and Hulme (1999) used a semiclosed pool of words in which item information played a more important role. They found an advantage for concrete over abstract words in both backward and forward recall. However, the forward and backward conditions were not run within the same experiment, and therefore the comparison between the two recall conditions should be made cautiously.

The Present Study

In four separate experiments, we manipulated the lexical variables that have been the most studied in the literature in the context of immediate serial recall, that is, the effects of word frequency, lexicality, semantic similarity, and word imageability. Participants
did both forward and backward recall. In the forward recall condition, participants had to write the words from left to right, from the first one in the list to the last one, in the same order they had been presented. In the backward recall condition, participants were required to write the words from right to left, starting with the last word of the list until the first one. Item and order errors were analyzed. According to the item and order trade-off hypothesis, the lexical variables should have no effect in backward recall, because they mostly influence item memory. Moreover, if less weight is attributed to the item dimension in backward recall, one should observe an increase in item errors in backward recall compared with forward recall.

**Experiment 1: Word Frequency**

In the first experiment, we manipulated word frequency. The word frequency effect refers to the finding that lists of high-frequency words are better recalled than are lists of low-frequency words (e.g., see Hulme et al., 1997, 2003; Poirier & Saint-Aubin, 1996; Roodenrys & Quinlan, 2000; Saint-Aubin & LeBlanc, 2005; Saint-Aubin & Poirier, 2005; M. J. Watkins et al., 2000). Half of the sequences consisted of six low-frequency words, and the other half consisted of six high-frequency words. After each list, participants were cued to recall the words either in forward or in backward order.

In Bireta et al.’s (2010) study, participants were not informed about the direction of recall until after list presentation in order to ensure that the same processes were involved during list presentation in both forward and backward recall. Surprenant et al. (2011) showed that knowledge of recall direction prior to list presentation did not influence the effect of backward recall on the word length effect. Manipulating lexical variables, however, might affect the strategies used during encoding. In order to test if the strategies used during list presentation differ between forward and backward recall when lexical variables are manipulated, two different experiments were done. In Experiment 1A, participants had no prior knowledge of recall direction until after list presentation, whereas in Experiment 1B, participants knew the recall direction before list presentation.

**Experiment 1A**

In this experiment, the forward and backward recall trials were presented randomly. Participants were not told the recall direction until after the whole list had been presented.

**Method.**

**Participants.** A total of 20 undergraduate students from the Université de Moncton volunteered to participate in this experiment.

**Materials.** The experiment was controlled by a PC computer using E-Prime 1.0 with a resolution of 800 × 600 pixels. The stimuli were lists of six French words that had three or four phonemes. Low-frequency words had a frequency ranging between 0.07 and 2.97 occurrences per million \((M = 1.66, SD = 0.75)\), and high-frequency words had a frequency ranging between 100.27 and 1,140.14 occurrences per million \((M = 265.71, SD = 198.95)\) according to Lexique (New, Pallier, Brysbaert, & Farrand, 2004). The 156 low-frequency words and the 156 high-frequency words were distributed quasirandomly in lists of six words with the restriction that the words within a list did not rhyme or have the same meaning. Of the 52 lists, 48 were used as experimental trials and four were used as practice trials. The words were presented in black, lowercase, 28-point Arial font, in the center of the computer screen.

**Design.** There were three repeated-measures factors: recall direction (two levels: forward, backward), frequency (two levels: high, low), and serial position (six levels: from 1 to 6). The 48 experimental lists were divided into two sets of 24 lists (12 high- and 12 low-frequency lists). Each of the two sets of 24 lists was presented in the forward condition for half of the participants and in the backward condition for the other half. The 12 sequences in each of the four conditions (forward low-frequency, backward low-frequency, forward high-frequency, backward high-frequency) were presented in a different random order for each participant. Within a list, too, the words were presented in a different random order for each participant, which was determined by the program at the beginning of the experiment. The 48 experimental trials were preceded by the four practice trials, one in each of the four conditions.

**Procedure.** Participants were tested individually within one session lasting about 25 min. Participants sat in front of a computer screen at about 60 cm from the screen, on which the to-be-remembered items were presented. In each trial, six words were presented sequentially at a rate of one word per second (1,000 ms on/0 ms off). After the presentation of the last word, participants were presented with the capitalized word NORMAL [normal] in blue or INVERSE [reverse] in red letters in the center of the screen, just above the symbol \(\Rightarrow\) or \(\Leftarrow\) presented in black ink. In the forward recall condition, participants had to recall, on the answer sheet provided, the list from left to right, from the first word to the last word that had been presented. In the backward recall condition, participants had to recall the list from right to left, from the last word to the first word of the list. They were instructed to write an 'X' when they could not recall a word at a given serial position. Participants were not allowed to backtrack in order to modify a previously written answer. Participants initiated the next trial by pressing the space bar. The experimenter remained in the room during the four practice trials to ensure compliance with the instructions.

**Results.** Responses were first scored with a strict serial recall criterion. Separate analyses were then performed on item and order errors. All repeated-measures analysis of variance carried out on the proportion of correct responses and on the proportion of item and order errors in the five experiments are reported in the Appendix. In all analyses, the .05 level of significance was adopted and the Greenhouse-Geisser correction was applied when the sphericity criterion was not met.

**Strict serial recall criterion.** To be scored as correct, a word had to be recalled in its correct serial position. All misspelled words that could be unambiguously attributed to a presented word were scored as correct. Figure 1 shows that the probability of correct recall was slightly lower when words were recalled in the forward direction \((M = .40, SD = .14)\) compared with the backward direction \((M = .48, SD = .15)\) and higher for high-frequency words \((M = .53, SD = .16)\) compared with low-frequency words \((M = .36, SD = .13)\). The frequency effect was smaller in backward recall than in forward recall, although simple main effects showed that it was significant for both forward, \(F(1, 19) =\)
MSE/H11005.00, MSE/H9257.00, p^2/H11005.83, and backward, F(1, 19)/H11005.51.98, MSE/H11005.00, MSE/H9257.00, p^2/H11005.87, as recall directions. The analysis also showed a significant interaction between recall direction and serial position, which indicates that the forward recall condition yielded a greater primacy than recency effect, whereas the backward condition showed a greater recency than primacy effect. Finally, the significant interaction between frequency and serial position suggests that the effect of frequency is greater for the first than for the last serial positions.

**Item errors.** An item error was defined as a missing item (when participants wrote an X) or a recalled word that was not part of the list. The proportion of item errors was computed in each of the four conditions by dividing the number of item errors by the number of items presented (see Poirier & Saint-Aubin, 1996). Figure 2 shows that there are more item errors in the forward recall direction (M = .50, SD = .13) than in the backward recall direction (M = .41, SD = .12) and more item errors for low-frequency words (M = .56, SD = .13) than for high-frequency words (M = .36, SD = .13). There was a significant interaction between recall direction and frequency. Simple main effects confirmed that there were more item errors for low- than for high-frequency words in both forward recall, F(1, 19) = 123.92, MSE = .00, p^2 = .87, and backward recall, F(1, 19) = 121.95, MSE = .00, p^2 = .87.

**Order errors.** An order error was calculated when an item was not recalled in its correct position. The proportion of order errors was computed by dividing the total number of order errors by the total number of items recalled in each trial, regardless of the order (Murdock, 1976; Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999a). As shown in Figure 2, neither frequency nor recall direction modulated the proportion of order errors. The proportion of order errors did not significantly differ between forward recall (M = .22, SD = .15) and backward recall (M = .18, SD = .11) or between high-frequency words (M = .20, SD = .12) and low-frequency words (M = .20, SD = .13).

**Experiment 1B**

In Experiment 1B, participants were informed of the recall direction before the list was presented. We replicated Experiment 1A using a design where the forward and backward trials were presented in two separate blocks. The forward and backward conditions were blocked to minimize any carryover effect from one recall condition to another. Participants received instructions relative to the recall condition and performed two practice trials before each block.

**Method.**

**Participants.** A total of 28 undergraduate students from the Université de Moncton volunteered to participate in this experiment. These participants did not take part in Experiment 1A.

**Materials.** The materials were the same as those used in Experiment 1A.

**Design and procedure.** A mixed design was used with one between-subjects factor, block order (two levels: forward–backward, backward–forward), and three within-subjects factors, recall direction (two levels: forward, backward), frequency (two levels: high, low), and serial position (six levels: from 1 to 6). The two sets of 24 lists each comprising 12 high- and 12 low-frequency lists were presented in two separate blocks. The instructions rela-

![Figure 1](image1.png)

**Figure 1.** Proportion of correct responses as a function of serial position and word frequency in the forward and backward recall directions in Experiments 1A and 1B. Error bars represent 95% confidence intervals.

![Figure 2](image2.png)

**Figure 2.** Proportion of item errors and order errors as a function of recall direction and word frequency in Experiments 1A and 1B. Error bars represent 95% confidence intervals.
tive to the recall procedure were given prior to each block. Two practice trials in each of the forward and backward conditions were done prior to the block in the corresponding condition. The procedure was the same as that for Experiment 1A except that after each trial, the word RAPPEL [recall] appeared to prompt recall. The word was written in blue capital letters above the black arrow $\Rightarrow$ in the forward recall block and in red capital letters above the black arrow $\Leftarrow$ in the backward recall block.

Results. There was no main effect of block order, and block order did not interact with any other variable. Therefore, the effect of block order is not reported in the analyses and is not discussed further.

Strict serial recall criterion. Figure 1 shows that the probability of correct recall did not differ between the forward ($M = .49, SD = .11$) and backward ($M = .49, SD = .11$) recall directions. The probability of correct recall was higher for high-frequency words ($M = .58, SD = .12$) than for low-frequency words ($M = .40, SD = .10$). The frequency effect was smaller for backward recall than for forward recall, and this reduction was probably more important for the first serial positions, as suggested by the significant three-way interaction. Simple main effects showed that the frequency effect was significant in both forward, $F(1, 27) = 84.41, MSE = .01, \eta^2_p = .76$, and backward, $F(1, 27) = 64.54, MSE = .01, \eta^2_p = .71$, recall directions. The significant interaction between recall direction and serial position indicates that the forward recall condition produced a stronger primacy effect and a weaker recency effect than did the backward recall condition.

Item errors. The proportion of item errors was not statistically different between the forward ($M = .44, SD = .10$) and backward ($M = .43, SD = .09$; see Figure 2) recall directions. There were more item errors for low-frequency words ($M = .54, SD = .10$) than for high-frequency words ($M = .32, SD = .09$). There was a significant interaction between recall direction and frequency. Simple main effects indicated that the effect of frequency was significant for both forward recall, $F(1, 27) = 168.23, MSE = .01, \eta^2_p = .86$, and backward recall, $F(1, 27) = 121.25, MSE = .00, \eta^2_p = .82$.

Order errors. As shown in Figure 2, the proportion of order errors did not differ between forward recall ($M = .14, SD = .08$) and backward recall ($M = .15, SD = .08$) or between high-frequency words ($M = .15, SD = .07$) and low-frequency words ($M = .14, SD = .08$). The interaction between recall direction and word frequency was not significant.

Discussion. Experiments 1A and 1B replicated the word frequency effect by showing that low-frequency words are more difficult to recall than high frequency words (e.g., see Hulme et al., 1997; Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 2005). In addition, item and order analyses replicated previous findings with an effect of word frequency on the recall of item information but not on that of order information (Hulme et al., 1997, 2003; Morin, Poirier, Fortin, & Hulme, 2006; Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 2005; Stuart & Hulme, 2000). As observed in previous studies, recall direction altered the shape of the serial position curve: Forward recall was characterized by a marked primacy and a weaker recency effect, whereas backward recall yielded a strong recency effect with a less marked primacy effect (e.g., see Bireta et al., 2010; Li & Lewandowsky, 1995). Importantly, our results show that the frequency effect was only slightly reduced in backward recall. Indeed, performance in both forward and backward recall was strongly influenced by word frequency. Contrarily to the predictions made by the item and order trade-off hypothesis (Bireta et al., 2010), backward recall had no effect (Experiment 1B) or decreased (Experiment 1A) the number of item errors compared with forward recall.

In line with the results of Surprenant et al. (2011), the same pattern of results was obtained whether participants had knowledge of recall direction prior to list presentation (Experiment 1B) or not (Experiment 1A). Moreover, when the recall conditions were presented in two different blocks (Experiment 1B), the same pattern of performance was obtained for forward and backward recall. These findings suggest that the same processes are involved during encoding for the forward and backward conditions. In the next experiments, the procedure of Experiment 1A, where the backward and forward trials were presented in a random order, was used for the sake of comparability with Bireta et al.’s (2010) study.

Finally, we also observed that the effect of word frequency was strong for initial serial positions and decreased thereafter. This finding might seem unusual, since some studies have shown that the word frequency effect increased over serial positions (e.g., see Hulme et al., 1997; see also Schweickert, Chen, & Poirier, 1999). A careful examination of the literature, however, shows that the interaction between word frequency and serial position varies greatly depending on the specific methodology employed. Indeed, Hulme et al. (1997) used a closed pool of items presented auditorily and showed that the effect of word frequency increased over serial positions. When a visual presentation is used, combined with an open pool of words, as used here, several studies have shown that the effect of word frequency decreased as a function of serial position (e.g., see Poirier & Saint-Aubin, 1996; Roedentrys & Quinlan, 2000, Experiment 2; O. C. Watkins & Watkins, 1977).

Experiment 2: Lexicality

In the second experiment, we manipulated lexicality. The effect of lexicality refers to the finding that words are easier to recall than nonwords (e.g., see Hulme et al., 1991; Multhaup et al., 1996; Saint-Aubin & Poirier, 2000). Half of the lists comprised nonwords that consisted of consonant–vowel–consonant trigrams. The other half of the lists comprised words that were equated in pronunciation times to lists of nonwords.

Method. Participants. A total of 20 undergraduate students from the Université de Moncton volunteered to take part in this experiment. None had participated in the previous experiments.

Materials. The stimuli were words and nonwords used by Saint-Aubin and Poirier (2000). The 120 nonwords were distributed quasi-randomly in 24 lists of five nonwords with four restrictions—the first three proposed by Hilgard (1951): (1) no letter appeared more than once in the same list, (2) the ending consonant of one nonword never immediately preceded the initial consonant of the following nonword in the alphabet, (3) nonwords were not real words in French or well-known abbreviations, and (4) nonwords within the same list did not rhyme. The words contained an average of three syllables and were selected because they had the same pronunciation time as nonwords (545 ms/word, 548 ms/ nonword; Richard, 1993).
Procedure and design. The procedure and design were the same as those used in Experiment 1A.

Results.

Strict serial recall criterion. Figure 3 shows that the probability of correct recall was lower in the forward recall condition ($M = .45, SD = .15$) compared with the backward recall condition ($M = .50, SD = .15$) and higher for words ($M = .58, SD = .18$) than for nonwords ($M = .36, SD = .12$). The lexicality effect did not vary as a function of recall direction and serial position. As indicated by the significant interaction between recall direction and serial position, the forward recall condition yielded a greater primacy than recency effect, whereas the backward condition showed a greater recency than primacy effect.

Item errors. Figure 4 shows that there were more item errors for nonwords ($M = .60, SD = .11$) than for words ($M = .36, SD = .14$). The proportion of item errors for the forward ($M = .49, SD = .11$) and backward ($M = .47, SD = .33$) directions did not differ significantly. The interaction between recall direction and lexicality was not significant.

Order errors. Figure 5 shows that there were more order errors for words ($M = .18, SD = .09$) than for nonwords ($M = .09, SD = .11$). The proportion of order errors did not differ between the forward ($M = .15, SD = .12$) and backward ($M = .12, SD = .07$) recall directions. The interaction between recall direction and lexicality was not significant.

Discussion. Experiment 2 replicated the lexicality effect by showing that lists of words were better recalled than lists of nonwords (e.g., see Hulme et al., 1991; Multhaup et al., 1996; Saint-Aubin & Poirier, 2000). This effect was due to a better memory for item information in the word condition compared with the nonword condition. We also found that order errors were slightly more frequent for words than for nonwords, a finding that is similar to that reported in previous studies (e.g., Fallon, Mak, Tehan, & Daly, 2005; Saint-Aubin & Poirier, 2000). Backward recall produced a stronger recency effect but a weaker primacy effect compared with forward recall (see Li & Lewandowsky, 1995). Unlike in Experiment 1, however, in Experiment 2 the lexicality effect was not modulated by recall direction. Moreover, the level of item and order errors did not differ between the backward and the forward recall conditions.

Experiment 3: Semantic Similarity

When a list includes words that belong to the same semantic category, recall performance is higher than when list items are from different semantic categories (e.g., see Poirier & Saint-Aubin, 1995; Saint-Aubin et al., 2005; Saint-Aubin & Poirier, 1999b; Tse, 2009). Semantic similarity was manipulated in the third experiment. In half of the trials, lists comprised words that were from the same semantic category. In the other half, lists were composed of nonrelated words.

Method.

Participants. Twenty undergraduate students from the Université de Moncton volunteered to participate in this experiment. None had participated in the previous experiments.

Materials. The stimuli were lists of six French words adapted from the materials used by Saint-Aubin and Poirier (1999b; Experiment 2). From their 28 lists of seven semantically similar words, we retained 24 lists for the experimental trials and two lists for the practice trials. In each list, one word was chosen randomly
F(1, 19) = 10.15, MSE = .00, \( \eta^2_p = .35 \).

**Results.** The probability of correct recall was lower in the forward recall direction \( (M = .50, SD = .16) \) than in the backward recall direction \( (M = .55, SD = .13) \) and higher for semantically similar words \( (M = .59, SD = .14) \) than for semantically dissimilar words \( (M = .46, SD = .14; \text{see Figure 3}) \). The significant interaction between recall direction and semantic similarity indicates that the semantic similarity effect was slightly smaller in the backward recall direction than in the forward recall direction. Simple main effects showed that the semantic similarity effect was significant for the forward recall direction, \( F(1, 19) = 59.60, MSE = .00, \eta^2_p = .76 \), and the backward recall direction, \( F(1, 19) = 31.24, MSE = .00, \eta^2_p = .62 \). The analysis also confirmed that recall direction modulated the shape of the serial position curve, with a greater primacy than recency effect in forward recall and a greater recency than primacy effect in backward recall. Finally, the significant interaction between semantic similarity and serial position suggests that the effect of semantic similarity is greater for the first than for the last serial position.

**Item errors.** The analysis of item errors showed that there were more item errors in the forward \( (M = .36, SD = .10) \) than in the backward \( (M = .31, SD = .09) \) recall direction and more item errors for semantically dissimilar \( (M = .42, SD = .10) \) than for semantically similar words \( (M = .25, SD = .09; \text{see Figure 4}) \). The interaction between recall direction and semantic similarity was not significant.

**Order errors.** Order errors did not vary significantly between forward \( (M = .24, SD = .15) \) and backward \( (M = .20, SD = .11) \) recall and did not vary between semantically similar \( (M = .23, SD = .12) \) and dissimilar \( (M = .21, SD = .14) \) words. However, the interaction between recall direction and semantic similarity was significant. An examination of Figure 5 suggests that this interaction is due to the slight decrease in the backward recall condition for semantically different words. Indeed, simple main effects confirmed that the proportion of order errors was similar between the semantically similar and dissimilar words in the forward recall direction \( (F < 1) \) but was higher in the similar than in the dissimilar lists for backward recall, \( F(1, 19) = 10.15, MSE = .00, \eta^2_p = .35 \).

**Discussion.** In line with the results of Poirier and Saint-Aubin (1995; see also Saint-Aubin & Poirier, 1999b), lists of semantically related items were better recalled than were lists of nonrelated items. This effect occurred because of a better memory for item information in the semantically similar condition compared with the dissimilar condition. Experiment 3 replicated the results of the two previous experiments as well as those of previous studies (e.g., see Li & Lewandowsky, 1995) by showing that forward recall is characterized by a strong primacy and a less-marked recency effect, whereas the reverse was found in backward recall. As in Experiment 1, the semantic similarity effect was slightly reduced in backward recall, although the effect remained
significant in both forward and backward recall conditions. Moreover, backward recall showed a small decrease in the proportion of order errors for dissimilar lists. This effect might be due to the fact that in that particular condition, where items were different, participants relied more heavily on the temporal dimension. However, this was not at the expense of item memory, since item errors were not affected by recall direction. It may simply be that because items were more distinctive, they were easier to recall in the correct order, as postulated by SIMPLE (Brown et al., 2007).

**Experiment 4: Imageability**

*Imageability* here was defined as the facility of a word to evoke a mental image. Imageability is often used interchangeably with concreteness because both terms refer to the availability of a visual representation in memory (Acheson, MacDonald, & Postle, 2010). When to-be-remembered items are concrete words that are easy to imagine, recall performance is higher compared with lists of abstract words, which are more difficult to imagine (e.g., see Miller & Roodenrys, 2009; Romani et al., 2008; Walker & Hulme, 1999). The effect of imageability was manipulated in Experiment 4. In half of the sequences the words had a high imageability value, whereas words in the other half of the lists had a low imageability value.

**Method.**

**Participants.** Twenty undergraduate students from the Université de Moncton volunteered to participate in this experiment. None had participated in the previous experiments.

**Materials.** The stimuli were lists of six French words that were selected from the Omnilex database according to their imageability value, varying on a scale from 1 to 7. The 156 low-imageability words had a mean value of 2.66 (SD = 0.26), and the 156 high-imageability words had a mean value of 5.56 (SD = 0.30). All words had three or four phonemes. Low- and high-imageability words were paired on frequency, with a mean frequency of 41.1 and 40.3 occurrences per million for low- and high-imageability words, respectively. Low- and high-imageability words were assigned quasi-randomly to lists of six words, with the restriction that the words within a list did not rhyme or have the same meaning. Twenty-four lists in each condition were used as experimental trials, and two lists in each condition were used as practice trials.

**Procedure and design.** The procedure and design were the same as in Experiment 1A.

**Results.**

**Strict serial recall criterion.** Although recall performance was slightly higher in the backward recall direction (M = .42, SD = .11) than in the forward recall direction (M = .39, SD = .10), this difference was not significant (see Figure 3). Recall performance was higher for high-imageability words (M = .44, SD = .11) than for low-imageability words (M = .38, SD = .09). The effect of imageability did not differ between backward and forward recall. The significant interaction between recall direction and serial position indicates that the primacy effect was greater than the recency effect for forward recall but that recency was stronger than primacy for backward recall. Finally, the interaction between imageability and serial position suggests that the effect of imageability was stronger for initial compared with final serial positions.

**Item errors.** Figure 4 shows that there were more item errors for forward recall (M = .52, SD = .08) than for backward recall (M = .47, SD = .09) and more item errors for low-imageability words (M = .54, SD = .09) than for high-imageability words (M = .45, SD = .10). The interaction between recall direction and imageability was not significant.

**Order errors.** Order errors did not significantly differ between forward recall (M = .19, SD = .11) and backward recall (M = .19, SD = .09) and did not differ between low-imageability words (M = .20, SD = .10) and high-imageability words (M = .18, SD = .09). The interaction between recall direction and imageability was not significant.

**Discussion.** In this experiment, we replicated the effect of imageability by showing that words easy to imagine are better recalled than are words that are difficult to imagine (e.g., see Majerus & Van der Linden, 2003; Miller & Roodenrys, 2009; Romani et al., 2008; Tse & Altarriba, 2007, 2009; Walker & Hulme, 1999). Moreover, the analysis of item and order errors revealed that imageability enhances item memory without affecting order memory. With regard to item memory, the current results replicate previous studies using an open word pool (Tse & Altarriba, 2007, 2009) and a closed or a semiclosed word pool (Miller & Roodenrys, 2009; Walker & Hulme, 1999) with a forward recall procedure. For order errors, previous studies using a closed word pool produced mixed results that can be taken as suggesting that imageability has a small effect on order recall under those circumstances. In effect, with a forward recall procedure, Miller and Roodenrys (2009) found an imageability effect in their second experiment, in which 120 participants took part, but not in their first experiment, which included 40 participants. Similarly, although there was a trend in all three experiments conducted by Walker and Hulme (1999), it was significant in only one of them: their backward recall experiment. When an open pool is used, as it is the case in the present study, however, order errors seem not to be affected by word concreteness (Tse & Altarriba, 2007, 2009). In light of these findings, it seems that imageability has a strong effect on item recall whether an open or a closed word pool is used, while it has a limited effect on order recall only with a closed word pool. Our results show that recall direction does not modulate the effect of imageability on the pattern of errors.

**General Discussion**

The objective of the current series of experiments was to examine the effect of recall direction on four lexical variables as well as on item and order errors in order to test the item and order trade-off hypothesis (e.g., see Bireta et al., 2010). In Experiments 1–4, we replicated the effects of word frequency (e.g., see Hulme et al., 1997; Morin et al., 2006), of lexicality (e.g., see Hulme et al., 1991; Multhaup et al., 1996; Saint-Aubin & Poirier, 2000), of semantic similarity (e.g., see Poirier & Saint-Aubin, 1995; Tse, 2009), and of word imageability (e.g., see Miller & Roodenrys, 2009; Romani et al., 2008). In line with previous studies, our analysis of errors showed that these effects were mainly attributable to item memory, with no effect on order memory (see Saint-Aubin & Poirier, 1999a; Walker & Hulme, 1999), except for the effect of lexicality, which slightly affected the proportion of order errors (see also Fallon et al., 2005; Saint-Aubin & Poirier, 2000). In all experiments, we also reproduced the typical effect of back-
ward recall on the shape of the serial position curve: Forward recall yielded a strong primacy effect and a weaker recency effect, whereas backward recall produced a marked recency effect and a weaker primacy effect (e.g., see Li & Lewandowsky, 1995). Moreover, recall performance was slightly higher in backward recall than in forward recall in Experiments 1A, 2, and 3. Although it was not significant, this tendency was also observed for the effect of imageability in Experiment 4. At first glance, the advantage for the backward recall condition may seem surprising. Intuitively, this condition appears harder. However, a careful examination of previous studies using a similar procedure (immediate serial recall with an open pool) revealed either no difference between the forward and backward conditions (Li & Lewandowsky, 1995, Experiment 2; Thomas et al., 2003) or an advantage for the backward procedure over the forward procedure (see Madigna, 1971).

**Backward Recall and the Lexical Variables**

In all four experiments, we observed strong lexical effects in both forward and backward recall. The effects of word frequency and of semantic similarity appeared to be slightly attenuated in backward recall compared with forward recall, but this attenuation was small. Indeed, the effect size for the word frequency effect decreased from .83 and .76 in forward recall to .73 and .71 in backward recall in Experiments 1A and 1B, respectively, and that of the semantic similarity effect decreased from .76 in forward recall to .62 in backward recall. The magnitude of the word frequency and of the semantic similarity effects were therefore compelling in backward recall, and the effects of lexicality and of word imageability were completely unaffected by backward recall.

The presence of all four lexical effects in backward recall is in line with the findings of Walker and Hulme (1999), which showed an effect of concreteness in backward recall, but is in contradiction to the findings of Hulme et al. (1997), which revealed that the effect of frequency was abolished in backward recall. Walker and Hulme interpreted the discrepancy between their results and those of Hulme et al. as evidence for different processes underlying the effects of word frequency and of concreteness. We suggest that this discrepancy is more likely to be due to the use of different word pools. Indeed, Hulme et al. used a closed pool where the same four lists (short high-frequency, long high-frequency, short low-frequency, long low-frequency) were used throughout the experiment, whereas Walker and Hulme used a semiclosed pool containing four sets of 16 words. Because word frequency mainly influences item recall, and because the contribution of item information covaries with pool size, it is not surprising that Roodenrys and Quinlan (2000) showed that the magnitude of the word frequency effect was reduced when a closed pool of words was used in comparison with an open pool of words. Since we showed that the effect of word frequency is slightly reduced with backward recall, the smaller word frequency effect of Hulme et al. in forward recall would simply have vanished in backward recall. When the serial recall paradigm is used with an open pool of words, the effect of backward recall appears to be consistent across the four lexical variables. Indeed, backward recall had a limited impact on the magnitude of the effects. Therefore, contrary to the conclusion of Walker and Hulme (1999), our study suggests that all lexical effects might well be sustained by the same processes.

The idea that all lexical variables affect the same processes is in line with the reconstruction hypothesis, which has been suggested to account for the role of long-term memory in immediate recall (see, e.g., Hulme et al., 1991; Saint-Aubin et al., 2005; Schweickert, 1993; Stuart & Hulme, 2009). Simply put, the reconstruction hypothesis states that, at recall, the phonological representations set up by list presentation would be degraded, either through decay or interference. Before being output as a response, degraded representations would undergo a reconstruction or cleaning process. The reconstruction process would be based on long-term knowledge of the to-be-remembered items. For instance, words would be better recalled than nonwords because, to all practical means, adequate long-term representations would be available for only words. Within this account, item recall would largely benefit from the reconstruction process, while order recall would be only marginally affected. Although the reconstruction hypothesis is silent about the impact of recall direction, we argue that within this framework, whether participants recall the lists in forward or backward direction should have no effect on the lexical effects. In effect, the same reconstruction processes should operate irrespective of recall direction. Although the reconstruction hypothesis explains why lexical effects are preserved in backward recall, the overall effect of backward recall on performance is beyond its scope.

**Explaining Backward Recall**

The main effect of backward recall on performance in our study was to modify the shape of the serial position curve (see also Li & Lewandowsky, 1995). The principle of output interference seems to be sufficient to account for the present findings (e.g., see Bireta et al., 2010; Li & Lewandowsky, 1995). Indeed, in forward recall, the items first recalled are the items from the beginning of the list. As the list items are recalled, the items not yet recalled become degraded, leading to a weaker recency effect. In backward recall, the primacy effect is weakened because the representations of the first list items are degraded while the last list items are recalled. Our results also indicate that the level of performance was generally slightly higher in backward recall compared with forward recall. Recall performance could be higher in backward recall because the overall level of feature overwriting (e.g., see Nairne, 1990) or loss of discriminability (e.g., see Brown et al., 2007) is lower than in forward recall. For instance, in backward recall, the first items recalled do not suffer from interference by the presentation of other list items between their presentation and their recall. In forward recall however, memory for the first list items is impaired by the presentation of the rest of the list before they can be recalled.

This single-process view of backward recall provides a parsimonious account of the present results. An integrative explanation, however, should also account for the abolition of the benchmark memory phenomena reported by Bireta et al. (2010). Indeed, why would the lexical effects but not benchmark memory phenomena survive backward recall? Bireta et al. (see also Surprenant et al., 2011) suggested that the effect of backward recall on the benchmark memory phenomena was due to a trade-off between the recall of item and order information: In backward recall, more weight—or attention—would be given to the order dimension and less to the item dimension, leading to the abolition of item-based
effects, such as the phonological similarity effect and the word length effect. There are two main findings in the present study that are inconsistent with this account. First, according to the item and order trade-off hypothesis, the effects attributable to item memory should be abolished in backward recall, because less attention is given to the item dimension. Clearly, this was not the case in the present study: We showed that four lexical effects that are well known to affect item memory (e.g., see Saint-Aubin & Poirier, 1999a) were all preserved in backward recall. Second, if backward recall increases the weight of the order dimension at the expense of the item dimension, the proportion of item errors should be higher in backward recall than in forward recall. Rather, we observed that backward recall improved item memory in three out of five experiments (Experiments 1A, 3, and 4) and had no effect in the other two (Experiments 1B and 2). The pattern of results reported in the present study is therefore incompatible with the item and order trade-off hypothesis as postulated by Bireta et al.

One possibility is that lexical variables are not represented along item dimensions such as phonology or word length. Bireta et al. (2010) did not specify what specific features were represented by the item dimension, and Brown et al. (2007) acknowledged that additional mechanisms beyond those postulated by SIMPLE would be necessary to account for the effect of lexical variables. Therefore, if lexical variables affect immediate serial recall performance through some redintegration process (e.g., see Hulme et al., 1991) distinct from the actual item dimensions, it could explain why lexical effects are maintained in backward recall, whereas benchmark memory phenomena are abolished.

Another possibility is that lexical and phonological variables are represented along different item dimensions. In addition to a dimension representing phonology, items could be represented along a dimension that contains “richer multidimensional semantic representations” (Brown et al., 2007, p. 569) that would account for the effect of lexical variables. In forward recall, items such as high-frequency or imageability words would require less attention at retrieval than would low-frequency or imageability words. Lexical factors might still affect performance in backward recall because, although less attention is available for the recall of item information, high-frequency or imageability words still require less attention than do low-frequency or imageability words. Alternatively, the relative weight of the semantic and phonological dimensions could depend on the recall direction: More weight could be given to the semantic dimension in backward recall, whereas more weight would be given to the phonological dimension in forward recall. The latter view is consistent with one theory suggested by Hulme et al. (1997). Indeed, based on the finding that backward recall did not modulate the effect of concreteness, Hulme et al. suggested that backward recall relied more heavily on the retrieval of a semantic code, which contrasts with the phonological code privileged in forward recall. In backward recall, the lexical effects would be preserved because of the increased reliance on the semantic dimension, whereas the benchmark memory phenomena would be abolished because less weight is given to the phonological dimension.

One reason why backward recall would rely on the semantic dimension instead of a phonological dimension is that semantic coding might be more resistant than phonological coding. In line with the idea that general principles govern memory, it has been shown that order information might be maintained in an analogous way in long-term semantic memory more so than in immediate memory (Surprenant & Neath, 2009). For instance, the typical serial position curve has been observed with the U.S. presidents (Healy & Parker, 2001; Roediger & Crowder, 1976), the Canadian prime ministers (Neath & Saint-Aubin, 2011), and verses of well-known hymns (Maylor, 2002). Semantic retrieval might therefore be privileged when the task requirements change or are perceived as more difficult. Although it seems appealing, such an account suggests that lexical effects should increase in backward recall, where the semantic code is privileged. This was not the case in the present study. In this regard, simulations might be useful to examine the effect of varying the relative weight of the item-semantic, item-phonological, and order dimensions in order to understand the precise interaction between recall direction and lexical variables.

In conclusion, the present study showed that four lexical effects were preserved in backward recall and that backward recall had no effect on the recall of order information but generally improved item memory. This is difficult to reconcile with the item and order trade-off explanation, which suggests that backward recall increases the weight given to the order dimension at the expense of item recall (e.g., Bireta et al., 2010). Rather, our results are consistent with the reconstruction hypothesis (Hulme et al., 1991) and suggest that the same reconstruction processes operate in forward and backward recall. Although further investigation is required to demystify the processes underlying backward recall, we suggest that a shift from phonological to semantic retrieval strategies (see Hulme et al., 1997) could explain the effect of backward recall on both lexical effects and benchmark memory phenomena.

References


Saint-Aubin, J., & Poirier, M. (2000). Immediate serial recall of words and...

(Appendix follows)
### Appendix

**Results of Experiments 1 to 4**

**Table A1**

*Analyses of Variance for the Proportion of Correct Recall (Strict Serial Recall Criterion) and for Item and Order Errors*

<table>
<thead>
<tr>
<th>Source</th>
<th>Frequency (Exp. 1A)</th>
<th>Frequency (Exp. 1B)</th>
<th>Lexicality (Exp. 2)</th>
<th>Semantic similarity (Exp. 3)</th>
<th>Imageability (Exp. 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>F</td>
<td>MSE</td>
<td>$\eta^2_p$</td>
<td>df</td>
</tr>
<tr>
<td><strong>Strict serial recall criterion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>1, 19</td>
<td>28.93 $^*$</td>
<td>.03</td>
<td>.60</td>
<td>1, 26</td>
</tr>
<tr>
<td>LV</td>
<td>1, 19</td>
<td>114.96 $^*$</td>
<td>.03</td>
<td>.86</td>
<td>1, 26</td>
</tr>
<tr>
<td>SP</td>
<td>5, 95</td>
<td>36.07 $^*$</td>
<td>.13</td>
<td>.66</td>
<td>5, 130</td>
</tr>
<tr>
<td>RD × LV</td>
<td>1, 19</td>
<td>4.56 $^*$</td>
<td>.02</td>
<td>.19</td>
<td>1, 26</td>
</tr>
<tr>
<td>RD × SP</td>
<td>5, 95</td>
<td>59.87 $^*$</td>
<td>.05</td>
<td>.76</td>
<td>5, 130</td>
</tr>
<tr>
<td>LV × SP</td>
<td>5, 95</td>
<td>9.38 $^*$</td>
<td>.04</td>
<td>.33</td>
<td>5, 130</td>
</tr>
<tr>
<td>RD × LV × SP</td>
<td>5, 95</td>
<td>1.21</td>
<td>.02</td>
<td>.06</td>
<td>5, 130</td>
</tr>
</tbody>
</table>

| **Item errors** | | | | | | | | | | | | | | | |
| RD     | 1, 19 | 44.54 $^*$ | .00 | .70 | 1, 26 | 0.48 | .01 | 0.02 | 1, 19 | 2.33 $^*$ | .00 | .11 | 1, 19 | 13.17 $^*$ | .00 | .41 | 1, 19 | 17.42 $^*$ | .00 | .48 |
| LV     | 1, 19 | 213.74 $^*$ | .00 | .92 | 1, 26 | 245.75 $^*$ | .01 | .90 | 1, 19 | 160.99 $^*$ | .01 | .89 | 1, 19 | 183.97 $^*$ | .00 | .91 | 1, 19 | 28.25 $^*$ | .01 | .60 |
| RD × LV | 1, 19 | 4.20 $^*$ | .00 | .18 | 1, 26 | 5.51 $^*$ | .00 | .18 | 1, 19 | 0.00 $^*$ | .00 | .00 | 1, 19 | 2.05 $^*$ | .00 | .10 | 1, 19 | 0.32 $^*$ | .00 | .02 |

| **Order errors** | | | | | | | | | | | | | | | |
| RD     | 1, 19 | 2.36 | .02 | .11 | 1, 26 | 0.16 | .01 | 0.01 | 1, 19 | 2.21 $^*$ | .01 | .10 | 1, 19 | 2.72 $^*$ | .01 | .13 | 1, 19 | 0.00 | .01 | .00 |
| LV     | 1, 19 | 0.00 | .01 | .00 | 1, 26 | 0.73 | .01 | 0.03 | 1, 19 | 14.39 $^*$ | .01 | .43 | 1, 19 | 1.30 | .01 | .06 | 1, 19 | 0.77 | .01 | .04 |
| RD × LV | 1, 19 | 0.95 | .01 | .05 | 1, 26 | 2.69 $^*$ | .01 | .09 | 1, 19 | 0.98 $^*$ | .01 | .05 | 1, 19 | 6.68 $^*$ | .00 | .26 | 1, 19 | 0.15 | .01 | .01 |

*Note. Exp. = experiment; MSE = mean square error; RD = recall direction; LV = lexical variable; SP = serial position.

$^*$ $p < .05$.  

Received December 21, 2010  
Revision received April 28, 2011  
Accepted June 16, 2011