Psicológica (2001), 22, 71-96.

Categorical and order information in free recall of action phrases

Johannes Engelkamp* and Hubert Zimmer

Saarland University, Germany

Starting from the generate-recognise model of free recall, we will be addressing two points. First, we will discuss whether the enactment effect, i.e. better memory of self-performed actions (SPTs) compared to watched experimenter-performed actions (EPT) and to verbal tasks (VT) is due to enhanced relational and/or item information of SPTs. Second, we will propose that at least two different types of relational information must be distinguished: categorical and episodic (order) information. We will demonstrate (a) that categorical relational information is used equally in SPTs, EPTs and VTs, (b) that episodic relational information is more efficient in pure EPT than in pure SPT conditions, (c) that this advantage is lost when the two encoding conditions are mixed, and (d) that episodic relational information can be effectively used with short but not with long lists. This variability of episodic relational information determines whether SPT shows higher memory performances than EPT. Finally, we conclude that the SPT effect is based on item information because the SPT advantage is observed even though relational information is not enhanced.

Key words: Enactment effect, item and relational information, generate-recognise models of free recall.

It has been known for more than 30 years that item and relational information enhance memory performance in list learning. The positive effect of item information was particularly shown in experiments that were stimulated by the levels of processing approach (e.g. Craik & Lockhart,

---

* The experiments reported in this paper were supported by the German Science Foundation under En 124/13. Correspondence concerning this article should be addressed to: Prof. Dr. Johannes Engelkamp, Department of Psychology, University of the Saarland, PO Box 151150, D-66041 Saarbrücken, Germany. E-mail: engelkamp@rz.uni-sb.de

1972; Craik & Tulving, 1975; Hyde & Jenkins, 1969). In these experiments, participants work through orienting tasks which focus on item information. For instance, they assess whether a word fits into a sentence frame or whether it has a certain number of syllables, etc. Because the words of the to-be-learned list are usually unrelated and each word has its specific orienting task, the task induces item-specific processes. It turns out that deep semantic orienting tasks lead to better recall and recognition than shallow surface-oriented tasks. Dissimilar to tasks that promote item processing, there are tasks which focus on relational processing, that is, on the associations among the words of a list. An instruction to enhance relational encoding is, for instance, the request to form a story from the words of a list, or the task to sort the items into categories. Such relational instructions also enhance memory performance (e.g. Mandler, 1967; Mandler, 1968). Relational encoding of words is also enhanced when lists are categorically structured rather than unstructured (e.g. Kintsch, 1970).

These findings, that item and relational information enhance memory performance, were embedded and differentiated in the generate-recognise model (Bahrick, 1970; Kintsch, 1970). The basic assumptions of the model were as follows: Free recall is based on the generation of items which are then assessed in a second step, as to whether they belong to the learning episode. It is also assumed that in a recognition test, only the latter process, namely recognition, takes place because the to-be-assessed items are presented externally during testing. Hence, there is no need for item generation.

Hunt and Einstein (1981) directly tested the assumptions of the generate-recognise model by applying the principle of encoding specificity to it in the following way: If free recall is based on item information and relational information as proposed by generate-recognise models, it should be enhanced if either of them is enhanced. On the other hand, if recognition memory is primarily based on item information, it should only be enhanced if item information is enhanced. The direct measurement of relational information in recall by organisational scores is possible when categorically structured lists are used. Hunt and Einstein (1981) tested these assumptions. They manipulated categorical list structure as well as the instructions either to focus on item information (by a pleasantness rating) or on relational information (by a sorting task). They observed, as expected, that increasing item information and relational information increased free recall. Increasing item information increased recognition memory. Increasing relational information increased the organisational scores. In order to measure relational information in free recall, they used the adjusted ratio of clustering (ARC) score according to Roencker, Thompson and Brown

(1971) which assess the frequency with which items from the same category are recalled in succession.

There are two interesting aspects of the generate-recognise model and its test in the paradigm of Hunt and Einstein (1981). First, the paradigm facilitates insight with regard to the generate-recognise model, as to whether specific memory effects, such as the generation effect (e.g. Begg & Snider, 1987; Hirshman & Bjork, 1988) or the bizarre imagery effect (e.g. Einstein & McDaniel, 1987), are based on item and/or relational information. Second, categorical information is often considered to be the prototype of relational information. In any case, it is the most frequently studied type of relational information, and the question is rarely asked whether categorical information is the only type of relational information.

We believe that at least two broad types of relational information should be distinguished: categorical-relational information which is based on long-term conceptual knowledge and episodic-relational information which is grounded on new associations generated in the actual episode. The characteristic feature of categorical-relational information is that it is based on pre-experimental long-term knowledge. The list items are connected via their conceptual representations (i.e. their meanings) within the conceptual knowledge of a person. These connections are used to relate list items to each other. A direct neighbourhood of the items in the learning list is not a precondition for this mechanism to become effective. The situation is different for episodic relational information. This information refers to the actual neighbourhood of items in the learning list. This neighbourhood is incidental and not determined by the meaning of the items. The neighbourhood may be, and usually is, different from experience to experience. This kind of episodic relational information has been studied recently within the frame work of item-order information (e.g. DeLosh & McDaniel, 1996; Mulligan, 1999; Serra & Nairne, 1993). Order information refers to the neighbourhood of items in a specific list presentation.

Using the background provided above, we will address two questions in this paper with regard to a specific memory effect: the enactment or SPT (for subject-performed task) effect. The effect consists in the fact that simple action phrases are better retained when the actions which they refer to are performed rather than when the phrases are only listened to (for more information see below).

(1) Is the SPT effect based on item and/or relational information?

(2) Does the answer depend on whether categorical-relational information or episodic (order) relational information is studied?

The experimental results will show:

(a) Encoding of categorical-relational information is independent of the type of encoding condition, it does not contribute to the SPT effect.

(b) Encoding of episodic-relational information, on the other hand, is dependent on the type of encoding condition, it modifies the SPT effect.

(c) Moreover, the use of relational information is modified by the specific learning and testing situation.

The SPT effect

Before we deal with these questions, we will briefly introduce the SPT effect. In a typical experiment, participants are presented lists of unrelated phrases of simple actions in an audio or visual manner, such as "put the hat on", "open the box", "bend the wire". Half of the participants are requested to read or listen to the phrases and to memorise them. The other half is requested to perform the actions upon reading or hearing them and to memorise them. Participants may be given the objects to perform with or perform the actions symbolically (without real objects). After the presentation of the list, participants are requested to recall all items they remember in any order, or they are given the items together with distracter items in a recognition memory test. Under these conditions, memory performance was consistently better for SPTs than for the verbal tasks (VTs). This holds true for free recall as well as for recognition memory, and it holds true under a wide variety of list lengths and presentation rate conditions (see Engelkamp, 1998, for a review). Hence, the SPT effect is a quite robust effect. We will now address the question of how item, and particularly relational information, influence this effect. We will begin with categorical-relational information.

Categorical-relational information: SPTs versus VTs

As mentioned above, categorical-relational information is the most frequently studied type of relational information. It is mostly manipulated by list structure. In order to study whether categorical information manipulated by list structure influences the SPT effect, Zimmer and Engelkamp (1989) used the experimental paradigm introduced by Hunt and Einstein (1981). They constructed lists of simple action phrases that were categorically structured, and presented them either in VTs or in SPTs. They used two types of list structures: episodic and taxonomic. In the episodically structured lists, the items belonged to six different everyday episodes such as garden work or cooking. In the taxonomically structured lists, the items were from eight taxonomic categories such as cleaning activities or activities which involved emptying something. After the presentation of a list, the participants free-recalled what they remembered. The ARC scores were computed from the free recall data. Zimmer and Engelkamp expected the usual SPT effect, namely better free recall after SPTs than after VTs. If relational encoding contributes to the SPT effect, ARC scores should be larger for SPTs than for VTs. The results are summarised in Figure 1.

While there was the usual SPT effect in free recall with episodic as well as with taxonomic lists, there was no difference in the ARC scores of VTs and SPTs with either list. According to this finding, categorical-relational information does not contribute to the SPT effect.

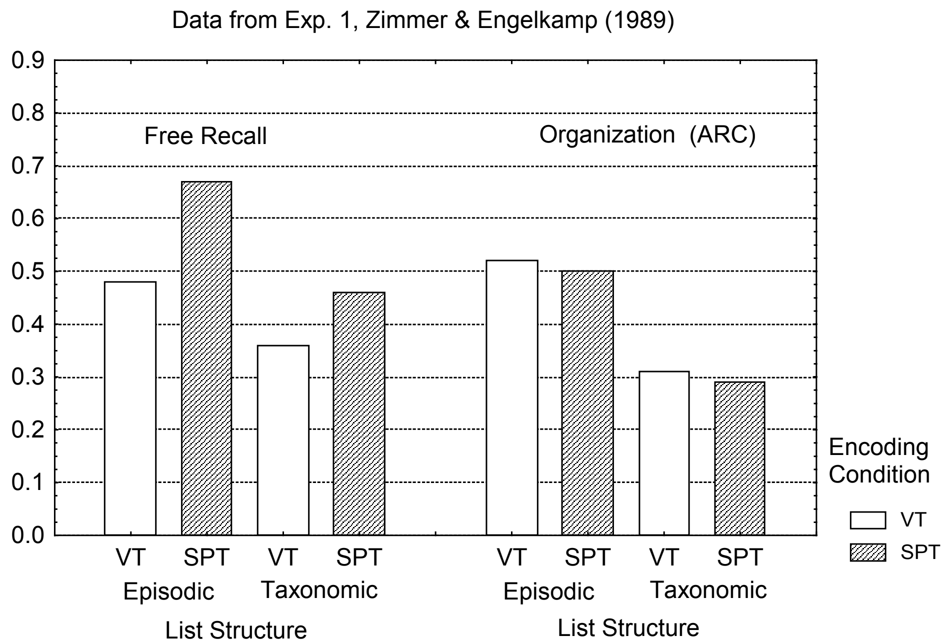

Figure 1. Mean relative frequencies of free recall and mean organization in recall (ARC scores) depending on list structure (episodic, taxonomic) and on encoding condition (verbal task = VT, subject-performed task = SPT) (after Zimmer, & Engelkamp, 1989, Exp. 1)

Unfortunately, this finding was not in agreement with results from a similar experiment by Bäckman, Nilsson and Chalom (1986). They used a list which was organised along categories of objects such as actions with toys, with clothes, etc. Furthermore, in addition to the usual learning conditions (VT, SPT), which were their control conditions, they used an interference condition. In this condition, participants had to achieve a continuous subtraction task during list learning. Their results are presented in Figure 2.

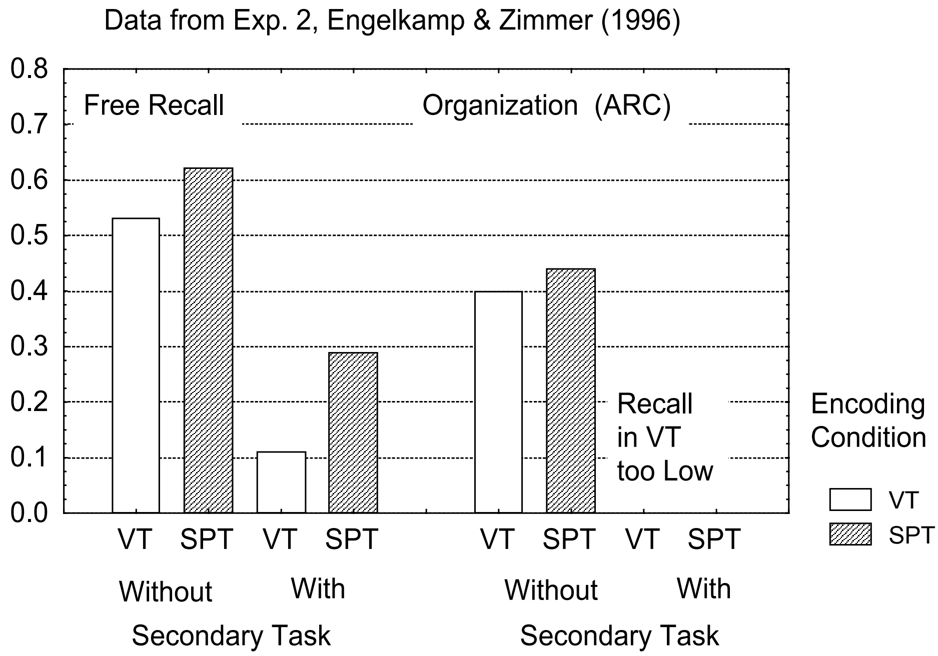

Figure 2. Mean relative frequencies of free recall and mean organization in recall (ARC scores) for a categorically structured list depending on a secondary task at encoding (with, without) and on encoding condition (verbal task = VT; subject-performed task = SPT) (after Bäckman, Nilsson & Chalom, 1986, Exp. 1).

They observed an SPT effect in free recall under control and under interference conditions. Moreover, they also observed better ARC scores after SPTs than after VTs. However, it cannot be excluded that the ARC advantage of SPTs in their experiment was due to the very low recall level and the negative ARC score when VTs were learned under interference conditions. In order to clarify the situation, Engelkamp and Zimmer (1996) replicated the experiment of Bäckman et al. (1986). Like Bäckman et al., they organised the lists according to the objects used and realised an interference in addition to a control condition. Deviating from Bäckman et al. in their experiments, a new/old recognition test was given after the free recall test. The results are summarised in Figure 3.

As expected, an SPT effect was found in free recall and also in the recognition test in Experiment 1 (VT.76, SPT.97). The recognition data are not presented in Figure 3. However, the ARC scores did not differ depending on experimental conditions. Because the ARC scores were small in this experiment, a floor effect cannot be excluded. The ARC scores were much lower than in the experiment from Bäckman et al. (1986) and from Zimmer and Engelkamp (1989). Engelkamp and Zimmer pointed to two possible reasons for the deviation of this finding: First, they used larger categories than Bäckman et al., and second, one of the two categories they used could have been thought of as a sub-category (beverages) of the other category (food). Therefore, they replicated the experiment with better controlled materials which were matched as closely as possible with the materials of Bäckman et al. (1986). They used five instead of four categories, made the categories more distinct, and added the category "body parts", because this category was used by Bäckman et al. (1986).

The results of this second experiment (except the recognition data) are also presented in Figure 3. As can be seen from Figure 3, Engelkamp and Zimmer (1996) succeeded in raising the ARC scores to the level of Bäckman et al. (1986). As expected, there were SPT effects in free recall and recognition memory (VT.60; SPT.82). However, there was still no ARC difference between VTs and SPTs. No ARC scores were computed for the interference condition because recall performance was only.11 with VTs.

After all, it seems likely that the findings from Bäckman et al. (1986) were due to their particular stimulus conditions. For instance, they might have used particularly salient objects. This feature could have been critical because in SPTs, they used real objects as compared to VTs in which no real objects were used. Unfortunately, their original list was no longer available when Engelkamp and Zimmer tried to replicate their results.

As a whole, the findings support the conclusion that relational encoding, based on categorical list structure, does not differ between VTs and SPTs. Hence, this relational information cannot be the cause for the SPT effect. The effect might rather be attributed to the better item information of SPTs compared to VTs. The conclusion that SPTs provide better item information than VTs is also supported by the SPT effect in recognition memory.

As to the influence of categorical list structure on SPT encoding, it might be argued that a categorical list structure, which is based on the taxonomic categories, is not the most appropriate or optimal way to test whether the SPT effect may also benefit from relational list structure information. Using script-like list structures might allow for a stricter test. However, Zimmer and Engelkamp (1989) used a list that was script-like (episodic list). They also observed, for this type of list, an SPT effect in free recall but no ARC advantage of SPTs. However, for their episodic list, one

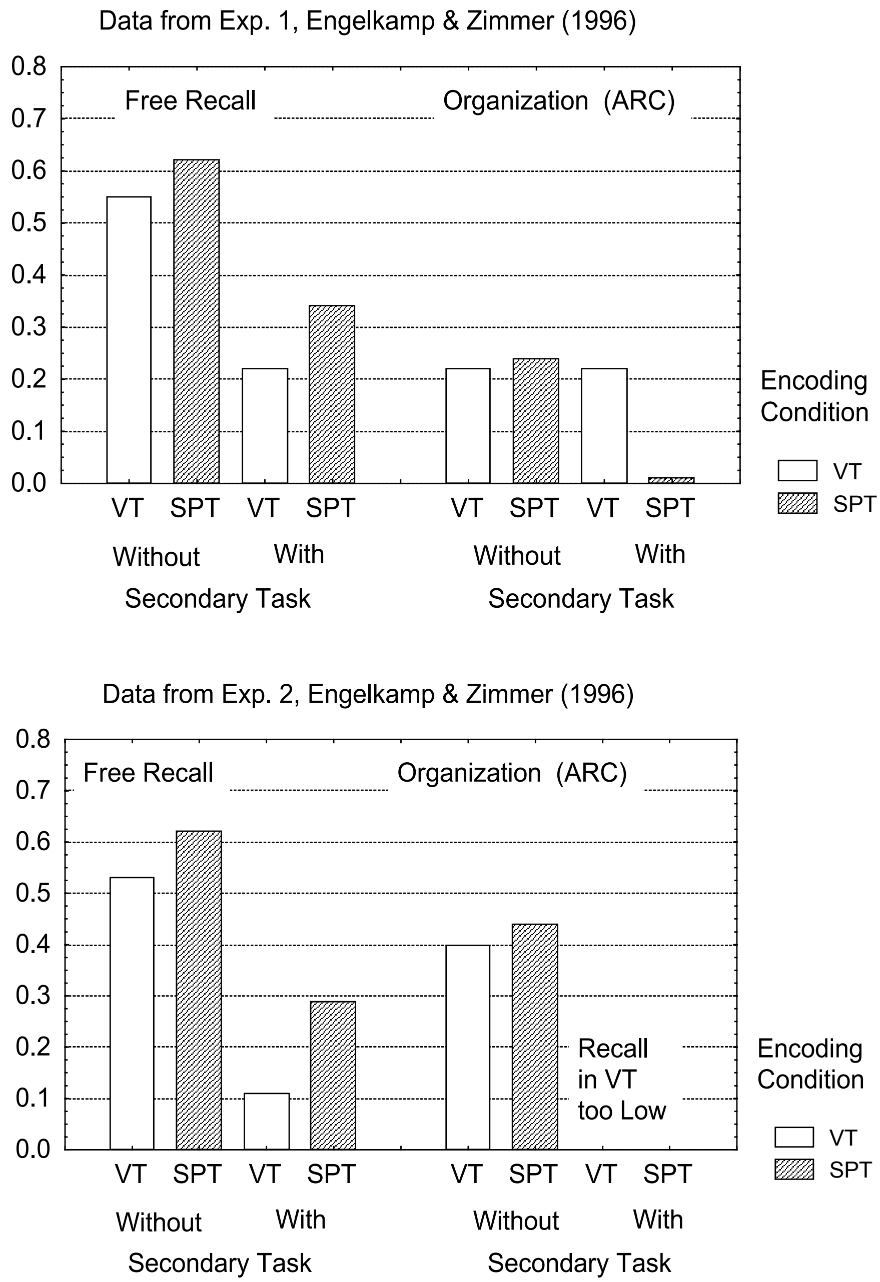

Figure 3. Mean relative frequencies of free recall and mean organization in recall (ARC scores) for a categorically structured list depending on a secondary task at encoding (with, without) and on encoding condition (verbal task = VT; subject-performed task = SPT) (after Engelkamp & Zimmer, 1996, Exp. 1 and Exp. 2)

might argue that this list was not an appropriate operationalisation of script information, because they used action verbs instead of action phrases.

Action verbs might have left the participants with too much lexical ambiguity because the meaning of verbs is often underdetermined without a given object context. For instance, verbs such as “cut” or “clean” might be used in many contexts (e.g. “cut wood” vs. “cut bread”, or “clean teeth” vs. “clean dishes”). They have different meanings in each context. Therefore, Engelkamp and Zimmer (2000) constructed script-like categories in using verb-object phrases. These script-like lists might be more useful in facilitating the activation of the corresponding script information with SPTs than with VTs.

In Experiment 1 from Engelkamp and Zimmer (2000), the list was presented to two groups of participants. The phrases were in random order. One group learned the list in SPTs, the other in VTs. The list consisted of 64 action phrases from eight script categories. Because the list was rather long, it was presented three times with a free recall after each presentation. The results are shown in Figure 4.

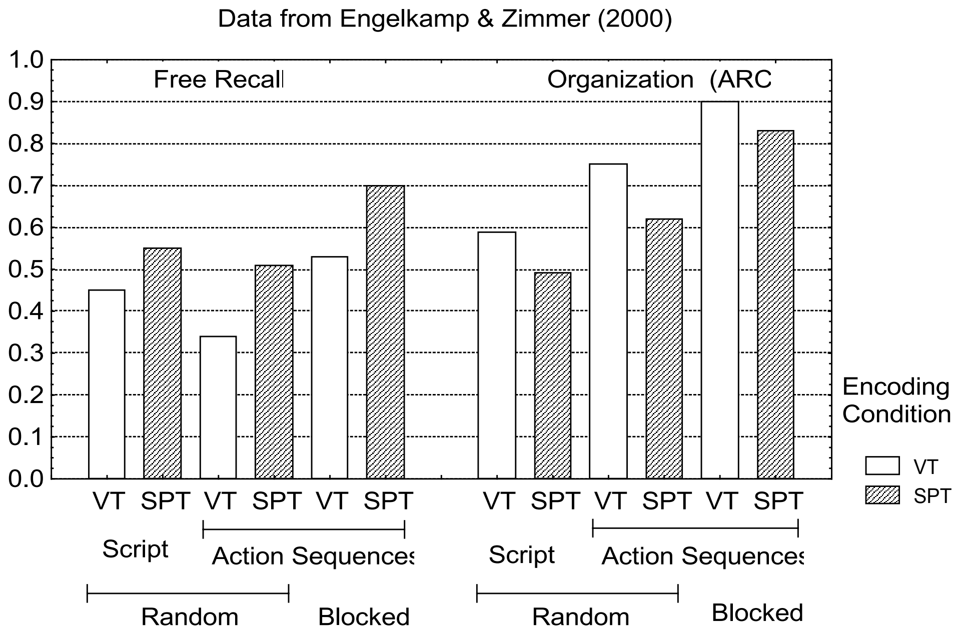

Figure 4. Mean relative frequencies of free recall and mean organization in recall (ARC scores) for script-like structured lists depending on type of encoding (verbal task = VT; subject-performed task = SPT). The list was presented randomly in Experiment 1 and randomly as well as blocked in Experiment 2 (after Engelkamp & Zimmer, 2000)

Again, there was a free recall advantage for SPTs over VTs and no statistically significant difference between SPTs and VTs for ARC scores. If anything, VTs showed a numerical advantage over SPTs. Both effects held true across trials. Although free recall, as well as organisational scores, increased over trials, there was no interaction of trials with the encoding condition.

The second experiment served as an even stricter test of the assumption that script information might be better used in SPTs than in VTs. In Experiment 1, the script information was presented randomly. Hence, the scripts were not immediately obvious from the list. In Experiment 2, the actions of a script were presented in their natural sequence. The items of each script were presented in one block in their natural order. For control purposes, the items were also presented in random order. In this experiment, the study material consisted of ten scripts with three actions each. An example of a script is car parking: stop the engine, take out the key, lock the door of the car. The results are presented in Figure 4. Also in this experiment, an SPT effect was observed, and again there were no significant differences between the ARC scores of SPTs and VTs. If anything, the organisation of VTs was slightly better than that of SPTs. Although free recall and ARC scores were greater with blocked than with random presentation, the factor blocked-random did not interact with the type of encoding.

The findings of the last two experiments show that script information is not used differently from taxonomic information. In both cases, there is no difference in the ARC scores between VTs and SPTs. If anything, in the last two experiments the ARC scores are smaller after SPTs than after VTs. In spite of the fact that there is no differential use of the script-like long-term information between SPTs and VTs as reflected in the ARC scores, there is an SPT effect in free recall. Therefore, the findings do not only show that the use of relational information in free recall, based on categorical list structure, does not differ between SPTs and VTs, but by showing this null-effect for relational information, they suggest at the same time that the SPT effect in free recall relies primarily on item information.

Categorical-relational information: SPTs versus EPTs

Actions cannot only be performed, the performance can also be observed by others. In the relevant experiments, the other is often the experimenter. Therefore, this task is called experimenter-performed task (EPT). The first question is whether there is also an SPT advantage over EPTs. As the literature shows, this is the case, although consistently only for long lists. We will later discuss the situation of short lists, which is more complex. Figure 5 gives a summary of some results.

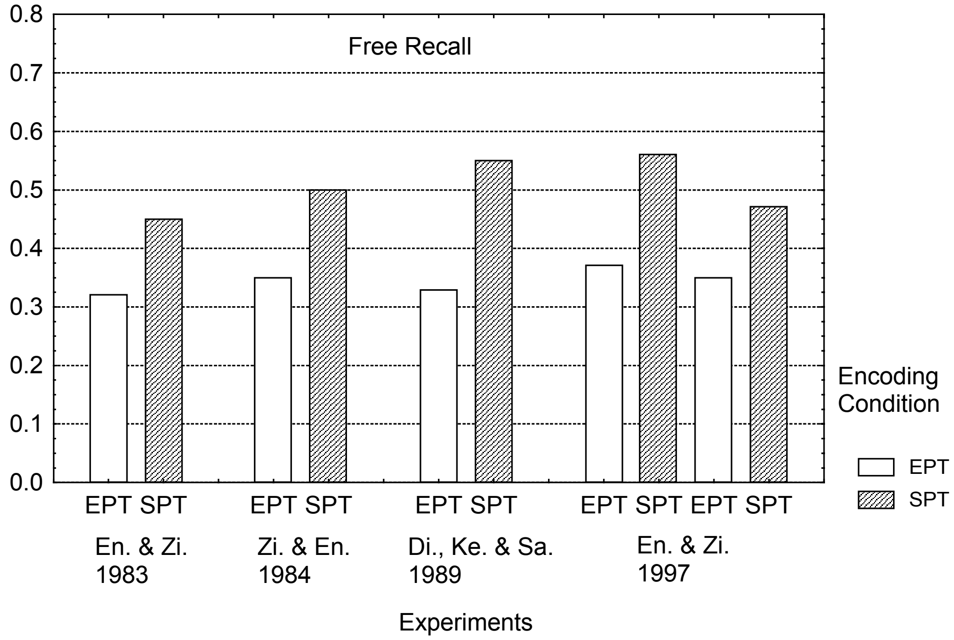

Figure 5. Mean relative frequencies of free recall as a function of type of encoding (experimenter-performed task = EPT, subject-performed task = SPT) in the following studies: Engelkamp & Zimmer, 1983; Zimmer & Engelkamp, 1984, Exp. 1; Dick, Kean and Sands, 1989; Engelkamp & Zimmer, 1997, Exp. 1 and Exp. 3

With long lists, there was consistently a better recall after SPTs than after EPTs. This SPT advantage was also observed when the participants imagined somebody performing the actions (e.g. Engelkamp & Krumnacker, 1980, Exp. 2).

Also with this SPT effect (SPT > EPT), the question arises as to whether the effect is based on item information, relational information or both. This question was tested in the same way as with the SPT advantage over VTs. Again, categorically structured lists were used because they allow for a direct measurement of relational information in computing ARC scores.

In a yet unpublished study, Engelkamp and Zimmer presented a list of 25 action phrases to two groups of 20 participants each. The list consisted of five categories with five exemplars each. The categories were determined by the object categories involved in the actions (such as toys, clothes, body parts, etc.). The action phrases were presented at a rate of 6 sec per item.

They were presented in random order. One group of participants listened to the phrases and observed the experimenter performing the actions (EPT). The other group self-performed the denoted actions (SPT). In both cases, the actions were performed symbolically (i.e. without real objects). After list presentation, there was an immediate free recall test (see Figure 6).

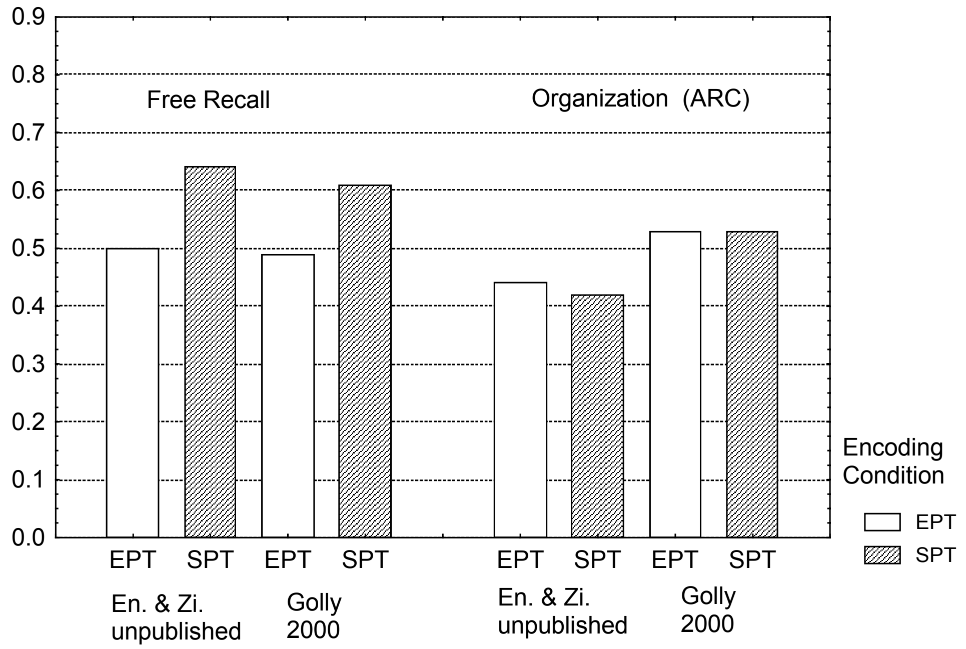

Figure 6. Mean relative frequencies of free recall and mean organization in recall (ARC scores) for categorically structured lists depending on type of encoding (experimenter-performed task = EPT, subject-performed task = SPT) (after experiments of Engelkamp & Zimmer, unpublished data, and of Golly, 2000)

Again, there was a clear-cut SPT effect in free recall but no difference in ARC scores between SPTs and EPTs. Relational information, as reflected in ARC scores, did not differ as a function of the type of encoding.

A similar study stems from Golly (2000). In her dissertation, she presented a categorically structured list consisting of 24 action phrases. The categories were structured by classes of objects such as food, clothes etc. There were four categories of six items in the list. One group of participants learned the list in EPTs and one group in SPTs. The results are also depicted in Figure 6. As the figure shows, free recall after SPTs was better than after EPTs. In spite of the SPT effect in recall, there was no difference in ARC scores between SPTs and EPTs. Golly also found that recognition defined as hits minus false alarms was better after SPTs than after EPTs. A recognition advantage of SPTs over EPTs was also reported by Saathoff (1999).

Taking the results as a whole, one can conclude that the use of relational information, based on categorically structured lists of action verbs and action phrases, as reflected in ARC scores in free recall performances, does not differ between tasks of encoding. In other words, categorical-relational information is used equally efficiently in VTs, EPTs and SPTs. In spite of this null effect, free recall and recognition are better after SPTs than after EPTs and VTs. This consistent pattern of findings shows that the SPT effect in free recall cannot be explained by better relational encoding of categorical information, be this information based on object categories, verb categories, episodes or scripts. Instead, it seems rather likely that the SPT effect in free recall is based on better item information of SPTs than of EPTs and VTs. However, this conclusion might be premature. It relies on the assumption that in free recall (a) there are only two types of information involved, namely item and relational information, and (b) that categorical information is the only type of relational information. Several arguments speak against categorical information as the only basis of the generation process in free recall.

There must be other relational information than categorical information. At closer inquiry, it seems likely that there may be other types of relational information than categorical information used in free recall which might differ depending on the type of encoding. Several arguments can be put forward in favour of the assumption that retrieval or generation processes in recall are not based solely on categorical information.

First, the use of categorical information is far from perfect. The ARC scores are, with .50 or less in randomly presented categorically structured lists, far away from a maximal value of 1.00. That means the participants use categorical information but not in a perfect manner. Hence, they are probably using also other retrieval mechanisms. This conclusion is also supported by the observation that the same proportion of free recall goes along with quite different degrees of organisation. For instance, a free recall level of about 50% can be observed with mean ARC scores ranging from about .20 to about .90.

Second, sometimes the correlations between ARC and free recall scores were computed. In these studies, a positive correlation was sometimes observed with VT learning. However, the ARC scores did not usually correlate with recall performance with SPTs (Zimmer & Engelkamp, 1989; Engelkamp & Zimmer, 1996; Engelkamp & Zimmer, 2000). Moreover, the differential correlations suggest that different free recall is based on different mechanisms in VTs and SPTs.

Third, because an SPT effect is also observed in unrelated lists in which no categorical information is offered to the participants by list structure, categorical information can hardly be the main basis for retrieval processes in these cases.

Before we move to another retrieval mechanism in recall, we will briefly dwell on item information. Although we do not know yet whether relational information other than categorical information contributes to the SPT effect in free recall, the SPT effect in recognition tests supports the assumption that better item information is provided by SPTs than by VTs and EPTs. It seems likely that SPTs also benefit from the better item information provided in free recall. That SPTs provide better item information than VTs is a widely accepted position among action memory researchers (e.g. Bäckman & Nilsson, 1985; Cohen, 1981; Knopf, 1991; Kormi-Nouri, 1995; Nyberg, 1993; Zimmer, 1991; see Engelkamp, 1998, for a review). However, there is little direct evidence for this assumption in studies of free recall except for the fact that there is a robust SPT effect. We know of only one study from Zimmer (1991) which was more specific. On the background of a generation-recognition model, Zimmer showed that more of the generated items were recognised after SPTs than after VTs. We will now turn to studies which address the interplay of item and relational information in a different way, and which deal with the role of order information in free recall.

Order-relational information in SPTs and EPTs

As mentioned in the introduction, the particular neighbourhood of items, in unrelated lists, is accidental. Do participants encode these accidental order relations? Yes, they do so. Moreover, order encoding depends on the type of the encoding task and on the experimental design used. In the early nineties, Nairne and Serra (Nairne, Riegler & Serra, 1991; Serra & Nairne, 1993) suggested an item-order hypothesis according to which, under some encoding conditions, order encoding is better than under other conditions, and the reverse holds true for item encoding. As a result, in free recall, there may be a trade-off so that performance is about equal in both conditions. However, the order advantage vanishes if both encoding conditions alternate from one item to the next. In contrast, item information is rather independent of such alternations of encoding conditions. In order to facilitate the understanding of these complex assumptions, we will show how Serra and Nairne (1993) used the item-order hypothesis to explain the generation effect. Words which were generated from word fragments are recalled better than the complete words which were read. However, this advantage is confined to mixed lists, i.e., some words had to be read, others had to be generated. If one group of participants is given a list to be read and another group a list to be generated (pure lists), the generation effect in free recall disappears (e.g. Begg & Sider, 1987; Slamecka & Katsaiti, 1987).

In order to explain this finding, Serra and Nairne (e.g. 1993) proposed that generating items provides better item information than reading and that, in contrast, reading provides better order information than generating items. However, the latter assumption is confined to pure lists, that is if reading and generation do not take place alternately. If both conditions are realised in the same list (mixed list condition), than the better order encoding of reading is impaired and reduced to the level of generation.

Serra and Nairne (1993) tested the differential availability of order information by presenting short mixed and pure (either read or generate) lists and requesting their participants to reconstruct the order of item presentation. They assumed that this test measured order encoding. For this task, they presented the items in a random order during testing, and the participants had to allocate them to the original presentation positions. As expected, order reconstruction (OR) was better for read than for generated items in pure lists, and it was equally low for both conditions in mixed lists.

On the basis of this finding, they explained that the free recall of generated items was greater than the free recall of read items in mixed lists because better item information was provided by generation than by reading. The disappearance of the generation effect in pure lists was explained by a trade-off between item and order information. Moreover, in the background was the general assumption that free recall is based on a serial output strategy. Due to the serial output strategy, an advantage for read items was expected in pure lists because they provided better order information than generated items. On the other hand, the fact that the item information of generated items is better than that of read items should compensate for the order disadvantage of generated items. Hence, free recall should not differ in pure lists.

Engelkamp and Zimmer (1997) have suggested the application of this item-order hypothesis to the memory for action phrases presented in SPTs and EPTs. They pointed to the fact that there were inconsistent findings with regard to the SPT advantage over EPTs in free recall. For instance, Cohen (1981, 1983; Cohen & Bean, 1983) did not observe any differences in free recall of SPTs and EPTs, whereas others (e.g. Dick et al., 1989; Engelkamp & Zimmer, 1983, see Figure 5) observed SPT effects. Engelkamp and Zimmer (1997) initially assumed that list length was decisive because Cohen used much shorter lists than the other authors. However, they found that there was another factor - type of design - involved besides list length. They observed that with long lists (20 items and more) there was always a SPT advantage over EPT. However, with short lists there was a SPT effect only if a within-subjects design was used in which participants learned a mixed list that consisted of EPTs and SPTs. If a between-subjects condition was realised in which participants learned either an SPT list or an EPT list (pure lists), there was no difference in free recall of both lists. It turned out that Cohen had consistently worked with short lists in a between-subjects design.

In order to explain this pattern of findings, Engelkamp and Zimmer (1997) suggested that with short lists participants would use a retrieval strategy based on serial information, that EPTs provided better order information than SPTs in pure lists but not in mixed lists, and that SPTs consistently provided better item information than EPTs. These assumptions are summarised in Table 1.

Table 1. Assumption of the Item-Order Hypothesis Applied to EPT and SPT in Short Lists

		Between design, pure lists	
Item information	EPT	<	SPT
Order information	EPT	>	SPT
		Within design, mixed lists	
Item information	EPT	<	SPT
Order information	EPT	=	SPT

Table 2. Expectations for Free Recall, Order Reconstruction, and Recognition in Experimenter-Performed Tasks (EPTs) and for Subject-Performed Tasks (SPTs) as a Function of Design Type (Between, Within) According to the Item-Order Hypothesis

		Between design, pure lists	
Expectations for free recall	EPT	=	SPT
Expectations for order reconstruction	EPT	>	SPT
Expectations for recognition	EPT	<	SPT
		Within design, mixed lists	
Expectations for free recall	EPT	<	SPT
Expectations for order reconstruction	EPT	=	SPT
Expectations for recognition	EPT	<	SPT

Engelkamp and Dehn (2000) tested this speculation and applied the item-order hypothesis to the learning of short lists of action phrases. In their experiments, the participants learned eight lists consisting of eight phrases each. The lists were either learned with pure SPTs and with pure EPTs in two groups of participants and with mixed lists of EPTs and SPTs in a third group. After the presentation of each list, memory was tested in a free recall test or in an order reconstruction test. In one experiment, a final recognition memory test followed at the end of all eight lists. The expectations for the three memory tests are summarized in Table 2.

For recognition memory, it was assumed that there should be an SPT effect independent of design type (pure/mixed) because recognition memory only depends on item information, and this information is consistently better with SPTs than EPTs. This effect was observed. For order reconstruction, an EPT advantage over SPTs was assumed for pure lists because EPTs would provide better order information than SPTs under this condition. However, because in mixed lists of EPTs and SPTs, order information of EPTs should be reduced to the level of SPTs, no difference in order construction between SPTs and EPTs was expected. The pattern of findings corresponded to these expectations. Finally, it was expected that in free recall there should be an SPT effect over EPTs in mixed lists due to better item information of SPTs and equal order information in EPTs and SPTs. On the other hand, this SPT advantage should be reduced or abolished in pure lists because EPTs provide better order information than EPTs in this condition. This expected interaction was also observed. The results are summarised in Figure 7.

So far, the findings coincide with the item-order hypothesis. It seems as if free recall of unrelated lists is based on a serial output strategy and as if order information is particularly good in EPTs of pure lists. Furthermore, the assumption is supported that under "usual" conditions, that is if order information between SPTs and EPTs do not differ, the SPT effect is due to the better item information of SPTs compared to EPTs. Again, the conclusion may be too hasty. Also, the item-order hypothesis relies on the assumption that in free recall there are only two types of information involved (here item and order information). Now relational information is identified with order information. Again there are arguments which speak against the assumption that free recall is based on only one type of relational information, which in this case would be serial order information.

There must be relational information other than order information

There are several arguments that can be put forward against the assumption of the item-order hypothesis that order information is the only basis of free recall.

If the hypothesis is assumed to be true, it must clearly be confined to unrelated lists. As it has been made clear in the preceding information, the studies, in which categorically structured lists were used, clearly showed that categorical information - as reflected in ARC scores – was used in free recall. Hence, the item-order hypothesis should be modified to: Free recall is based on order information if no categorical information is provided. If this constraint is true, using categorically structured lists should make the interaction of type of encoding (EPT, SPT) and type of design (between, within) in free recall disappear. Only a main effect of type of encoding for categorically organised lists should be observed. It was demonstrated by Golly (2000) that this assumption holds true. She conducted an experiment using eight lists of eight items. The items of one list were always from one category. She presented the lists in random order and tested free recall and order reconstruction after each list. The analyses showed that there was a main effect of the type of encoding in free recall and no interaction. Free recall of SPTs (.84) was better than that of EPTs (.77) independently of the type of design. For order reconstruction, no effects were observed at all. The fact that retrieval strategies, other than the order-based one, are used if lists are categorised, was also shown by Serra and Nairne, Riegler and Serra (1991) and by Mulligan (1999) for other encoding conditions.

These findings support the assumption that the item-order hypothesis must be confined to unrelated lists. If a categorical list structure is presented, this structure is used during the encoding and retrieval of the items. Because categorical information is equally available with SPTs and EPTs, the advantage of item information of SPTs compared to EPTs determined the relative free recall performance. The fact that categorical information masks the order information can also be seen from the finding that order reconstruction scores do not differ any longer. The fact that categorical information overlays order information seems to be functional, given the superior recall performance of categorical lists compared to that of unrelated lists.

Another constraining factor of the item-order hypothesis is list length. It does not seem plausible that participants base free recall of long lists with 20 or more items on order information and on a serial output strategy. Rather it seems likely that with long lists participants resort to other retrieval strategies. Engelkamp and Dehn (2000) tested this assumption by using lists of 24 unrelated action phrases. They presented each participant four lists, two lists were always tested for free recall and two lists either for order reconstruction or for recognition. They observed the following results. First of all, order reconstruction scores were very low (<.15) and did not differ among the experimental conditions.

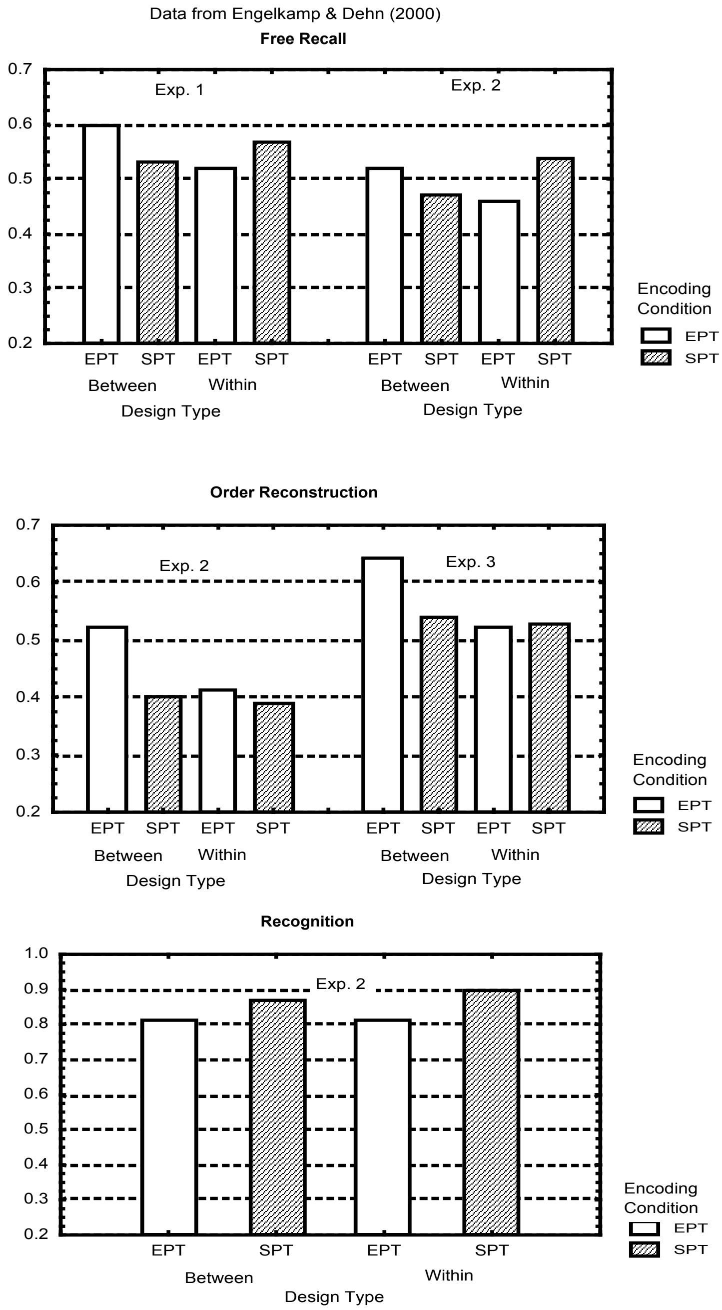

Figure 7. Mean relative frequencies of recognition (hits minus false alarms), order reconstruction, and free recall for short lists (8 items) depending on type of encoding (experimenter-performed task = EPT, subject-performed task = SPT) in different experiments of Engelkamp and Dehn (2000)

It should be noted that order reconstruction was not based here on absolute correct position but on relative correct position. A positive score was given when item X was correctly placed before item Y.

This more lenient score was used because otherwise the scores would have been even lower than.15. This finding makes it unlikely that this poor memory for order information is used as the basis for free recall. Free recall data supported the assumption that with long lists order-based retrieval is not an important determinant of recall. There was a main effect of type of encoding, but no interaction with the factor type of design. Recall of SPTs (.40 and.38) was better than of EPTs (.34). It is likely that the better free recall of SPTs than that of EPTs is due to better item information. This assumption corresponds with the main effect of the type of encoding in recognition (.94 for SPTs and.84 for EPTs) which was, as expected, independent of design type.

Relative order reconstruction scores were also computed for the first quarter of the lists. Here, the pattern was different from that for the whole list. In the between-subjects condition, the corresponding score was.33 for EPTs and.40 for SPTs, however, only.18 for the mixed list condition. This pattern suggests two things. First, that the participants might have used a retrieval strategy which was based on order information for the first items of the list, because the scores for order information were remarkably higher for the first quarter of the list than for the whole list. However, they must have used other retrieval strategies for the other three quarters of the list. Second, in mixed lists order information must have played a more minor role in free recall than in pure lists for the following reason: Although order information for mixed lists was lower than for pure lists, free recall performance did not differ between pure and mixed lists.

Hence, the findings from the long lists suggest that free recall is determined here, only partly by order information if at all. Most likely, participants may start with an order-based retrieval strategy and than move to other – as yet unknown – retrieval strategies. Moreover, the findings suggest that retrieval strategies differ for pure and mixed lists. The use of order-based retrieval strategies is less likely for mixed than for pure lists even for the first quarter.

The interesting assumption that participants use different strategies in recalling one and the same list was tackled further in another still unpublished experiment in our laboratory. The logic behind this experiment was as follows: If free recall of short lists of unrelated items was based indeed on an order-based serial output strategy, the explicit instruction to learn the list for a serial recall (SR) should hardly change the pattern of findings which were found when the instruction was given to learn the list just for recall. The free recall pattern should not differ between both instructions since participants encode order information and use it anyway in free recall.

The same lists of eight items as in Experiment 1 of Engelkamp and Dehn (2000) were presented using the same design with the type of encoding and type of design as factors and the same procedure. The only modification was that now the participants were explicitly requested to pay attention to the order of items and be prepared for a serial recall. The findings are presented in Table 3.

Table 3. Probability of Free Recall and Serial Recall as a Function of Type of Encoding (EPT, SPT), Type of Design (Between, Within) and Type of Instruction (FR Instruction, SR Instruction). The Data of FR Instruction Stem From Experiment 1 of Engelkamp and Dehn (2000).

	FR instruction		**SR instruction**	
Free recall	EPT	SPT	EPT	SPT
Between	.60	.53	.60	.57
Within	.52	.57	.53	.59
Serial recall				
Between	.07	.03	.25	.16
Within	.04		.14	

In order to facilitate the comparison between the effects of the two types of instruction (free versus serial recall), the free recall data of the experiment from Engelkamp and Dehn (2000) are presented again in Table 3. The scores for free recall contain all correctly produced items independently of whether they were recalled at the correct position or in the correct order. In serial recall, only those items were counted as correct which were also remembered in the correct order starting from the first position. Table 3 shows that there was practically no difference in either the recall level or the recall pattern as a function of the type of recall instruction if performance was scored for free recall. On the contrary, if performance was scored for serial recall, level of performance was clearly higher after a serial recall instruction than after a free recall instruction. However, the pattern did not differ. In both cases, performances under EPT-pure list were better than under the other conditions.

First of all, these findings show that the serial recall instruction was efficient. Participants under serial recall instructions showed a better serial recall than those under free recall instructions. The findings further confirm that serial order is best encoded in EPTs with pure lists. However, the findings also show that free recall performance does not depend on the type of instruction. This latter finding together with the serial recall findings suggests that the participants recall a substantial amount in addition to those items which they recall serially. It is very unlikely that these additional items were based on a serial retrieval strategy and that they represent simply errors in doing so correctly. It is much more likely that the participants use other retrieval strategies if their capacity for a serial retrieval is exhausted. This assumption is supported by two other findings. First, although serial recall is worse after a free than after a serial recall instruction, free recall performance does not differ. Second, although in EPTs with pure lists, order information is better available than in the other conditions, free recall of EPT-pure lists is hardly better than that of the other conditions, particularly of SPT in mixed lists.

As a whole, the experiments conducted in order to test the item-order hypothesis in the context of learning action phrases, have shown that the authors who pursued this hypothesis (e.g. Serra & Nairne, 1993; DeLosh & McDaniel, 1996) pointed to the important aspect that free recall may be based on a serial retrieval strategy, and that some encoding conditions allow for better order information encoding than others. However, they generalised the item-order hypothesis further than justified. Most importantly, they did not confine the hypothesis to short lists, and they did not recognise that order information is only one type of relational information on which retrieval in free recall may be based (cf. also McDaniel, DeLosh & Merritt, 2000, for a similar conclusion). The fact that the level of free recall can be much higher than the corresponding indices of order-based retrieval strategies, suggests that in free recall different types of retrieval mechanisms are used simultaneously.

CONCLUDING DISCUSSION

We started from the distinction between item and relational information and from the generate-recognise theories which assume that in free recall items are first generated and than assessed as to whether they are list items or not. These theoretical considerations leave open what is precisely meant by item and relational information. We focused primarily on relational information and its contribution to the recall of action phrases under different encoding conditions. We started with categorical information as one important type of relational information and later moved to order information as another type of relational information. We demonstrated that categorical information does not differ with action phrases dependent on the type of encoding task. We also demonstrated that order information depends on the type of encoding task under certain boundary conditions. However, we did not discuss in detail why these two types of relational encoding differ. We will do so here.

The types of categorical information that were addressed in this article – object-based taxonomies, action-based taxonomies, categories based on episodes, categories based on scripts – have one common feature : It refers to pre-experimental long-term semantic knowledge as shared by many persons. During encoding, the items of a list activate this long-term knowledge – e.g. the category of clothes or the category of garden work etc., and the activation of these long-term knowledge categories activate in turn their exemplars. During testing, the categories are initially retrieved because they have the strongest traces, and from the categories, their exemplars are activated and tested for their connection to the list episode. If one considers the possibilities of such a retrieval, it becomes plausible that each remembered item may also activate other items associated with it. These associations can, but must not belong to the categories as determined by the list. They can also be based on subjective associations which occurred during encoding. The items "open the umbrella" and "lift the coin" may, for instance, evoke the thought, in a particular person, that he or she needs a new umbrella and that it should not be too expensive. In short, categories, as they are used in categorically structured lists, are likely to activate long-term knowledge of these categories. However, other activation of associated long-term knowledge which is more idiosyncratic is also likely. Therefore, it is likely that categorical information is used which has been offered by the list. However, it is also likely that other long-term knowledge which is more idiosyncratic is used as well. The latter might be more important in unrelated than in related lists. We assume that this long-term knowledge is activated by encoding the action phrases which are a constant feature of all experimental tasks. What is added by the specific type of encoding task is more likely item information.

In sharp contrast to the use of any kind of long-term knowledge, is the relational order information. The specific order of items in an unrelated list has no basis in long-term knowledge. On the contrary, this is almost pure episodic information which changes from one list presentation to the next. Order information in a list is accidental, and we need a different mechanism to encode this information. In some theories, a special buffer is assumed which associates the items that are in it simultaneously. These are typically items presented as neighbours during encoding (e.g. Raijmakers & Shiffrin, 1981). However, these theories do not specify the role of encoding conditions in this association process. We assume that EPTs or VTs, in which the focus is on perceptual processes, are more appropriate to build new associations between neighbouring items than SPTs which focus on motor processes and thereby on each single item (see Engelkamp, 1995).The fact that SPTs force participants to focus on each action in order to be able to perform it smoothly, has two side effects: First, the focus on performing the actions increases item encoding, and second, the very same mechanism hinders order encoding. If SPTs intervene between EPTs in mixed lists, they also hinder the typical associations of EPTs because EPTs are put apart by SPTs, and SPTs interrupt the associative process. It also seems likely that although order information is encoded if items are simultaneously in the buffer, other encoding processes take place as well.

With these two briefly sketched types of relational information, one of which is independent of the type of encoding and the other which is not, we have identified two types of relational information which are obviously involved in list learning and which can be used in free recall. It is likely that other types of relational information are also involved in free recall. What they will be remains, for the time being, an unanswered question.

RESUMEN

Partiendo del modelo de recuerdo libre “generación-reconocimiento”, abordamos dos puntos. En primer lugar discutiremos si el efecto de “actuación”, es decir, el recuerdo superior de acciones que uno mismo ha realizado (AUR) sobre las acciones realizadas por otros (ARO) o sobre las tareas verbales (TV) se debe al fortalecimiento de la información relacional y/o de item en las condiciones de AUR. En segundo lugar, proponemos que es necesario distinguir entre, al menos, dos tipos de información relacional: categórica y episódica (de orden). Mostraremos a) que la información categórica se utiliza por igual en AUR, ARO y TV, b) que la información episódica relacional es más eficaz en la condición ARO pura que en la AUR. c) que esta ventaja se pierde cuando las dos condiciones de codificación se mezclan y d) que la información episódica relacional se puede utilizar de forma efectiva con listas cortas , pero no con listas largas. Esta variabilidad de la información episódica relacional determina que la condición AUR produzca mejor recuerdo que la ARO.

Palabras clave: Efecto de actuación, información de ítem e información relacional, modelo de generación-reconocimiento en recuerdo libre.

REFERENCES

Bäckman, L. & Nilsson, L.-G. (1985). Prerequisites for lack of age differences in memory performance. *Experimental Aging Research, 11*, 67-73.

Bäckman, L., Nilsson, L.-G. and Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition, 14*, 339-346.

Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review, 77*, 215-222.

Begg, I. & Snider, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 553-563.

Cohen, R.L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology, 22*, 267-281.

Cohen, R.L. (1983). The effect of encoding variables on the free recall of words and action events. *Memory & Cognition, 11*, 575-582.

Cohen, R.L. & Bean, G. (1983). Memory in educable mentally retarded adults: Deficit in subject or experimenter. *Intelligence, 7*, 287-298.

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

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268-294.

DeLosh, E.L. & McDaniel, M. (1996). The role of order information in free recall: Application to the word frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1136-1146.

Dick, M.B., Kean, M.L. & Sands, D. (1989). Memory for action events in Alzheimer-type dementia: Further evidence of an encoding failure. *Brain and Cognition, 9*, 71-87.

Einstein, G.O. & McDaniel, M.A. (1987). Distinctiveness and the mnemonic benefits of bizarre imagery. In M.A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes*. New York: Springer.

Engelkamp, J. (1995). Visual imagery and enactment in memory of actions. *British Journal of Psychology, 86*, 227-240.

Engelkamp, J. (1998). *Memory for actions*. Hove: Psychology Press.

Engelkamp, J. & Dehn, D. (2000). Item and order information in subject-performed tasks and in experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition* (in press).

Engelkamp, J. & Krumnacker, H. (1980). Imaginale und motorische Prozesse beim Behalten verbalen Materials. *Zeitschrift für experimentelle und angewandte Psychologie*, *27*, 511-533.

Engelkamp, J. & Zimmer, H. D. (1983). Zum Einfluß von Wahrnehmen und Tun auf das Behalten von Verb-Objekt-Phrasen. *Sprache & Kognition*, *2*, 117-127.

Engelkamp, J. & Zimmer, H. D. (1996). Organization and recall in verbal tasks and in subject-performed tasks. *European Journal of Cognitive Psychology, 8*, 257-273.

Engelkamp, J. & Zimmer, H. D. (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica, 96*, 43-60.

Engelkamp, J. & Zimmer, H.D. (2000). Taxonomic and script information in lists of action phrases do not influence the enactment in free recall. Submitted.

Golly, C. (2000). *Organisationsprozesse im Handlungsgedächtnis: Die Rolle relationaler Information für das Erinnern eigener und beobachteter Handlungen.* Unpublished Dissertation, Universität des Saarlandes, Saarbruecken.

Hirshman, E. & Bjork, R.A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 484-494.

Hunt, R.R. & Einstein, G.O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20*, 497-514.

Hyde, T.S. & Jenkins, J.J. (1969). The differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology, 82*, 472-481.

Kintsch, W. (1970). Models for free recall and recognition. In D.A. Norman (Ed.), *Models of human memory*. New York: Academic Press.

Knopf, M. (1991). Having shaved a kiwi fruit: Memory of unfamiliar subject-performed actions. *Psychological Research, 53*, 203-211.

Kormi-Nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology, 7*, 337-363.

Mandler, G. (1967). Organization and memory. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press.

Mandler, G. (1968). Association and organization: Facts, fancies, and theories. In T.R. Dixon & D.L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, NJ: Prentice Hall.

McDaniel, M., DeLosh, E.L. & Merritt, P.S. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition* (in press).

Mulligan, N.W. (1999). The effect of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 54-69.

Nairne, J.S., Riegler, G.L. & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 702-709.

Nyberg, L. (1993). *The enactment effect: Studies of memory phenomenon*. Unpublished doctoral dissertation, University of Umeå.

Raaijmakers, J.G.W. & Shiffrin, R.M. (1981). Search of associative memory. *Psychological Review, 88*, 93-134.

Roencker, D.L., Thompson, C.P. & Brown, S.C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76*, 45-48.

Saathoff, J. (1999). *Visuelle und motorische Information beim Erinnern von Handlungen.* Lengerich: Pabst.

Serra, M. & Nairne, J.S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. Memory & Cognition, 21, 34-40.

Slamecka, N.J. & Katsaiti, L.T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language, 26,* 589-607.

Zimmer, H.D. (1991). Memory after motoric encoding in a generation-recognition model. *Psychological Research, 53*, 226-231.

Zimmer, H.D. & Engelkamp, J. (1989). Does motor encoding enhance relational information? *Psychological Research, 51*, 158-167.