In *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum, 1995.

# Implicit Learning in the Presence of Multiple Cues

**Axel Cleeremans**[3]
Laboratoire de Psychologie Industrielle et Commerciale
Université Libre de Bruxelles
Avenue F.-D. Roosevelt, 50 — CP 122
1050 Bruxelles — BELGIUM
axcleer@ulb.ac.be

### Abstract

Is implicit learning an independent and automatic process? In this paper, I attempt to answer this question by exploring whether implicit learning occurs even despite the availability of more reliable explicit information about the material to be learnt. I report on a series of experiments during which subjects performed a sequential choice reaction task. On each trial subjects were exposed to a stimulus and to a cue of varying validity which, when valid, indicated where the next stimulus would appear. Subjects could therefore optimize their performance either by implicitly encoding the sequential constraints contained in the material or by explicitly relying on the information conveyed by the cue. Some theories predict that implicit learning does not rely on the same processing resources as involved in explicit learning. Such theories would thus predict that sensitivity to sequential constraints should not be affected by the presence of reliable explicit information about sequence structure. Other theories, by contrast, would predict that implicit learning would not occur in such cases. The results suggest that the former theories are correct. I also describe preliminary simulation work meant to enable the implications of these contrasting theories to be explored.

## Introduction

Implicit learning is typically defined as the process whereby subjects appear capable of acquiring new information without concomitant awareness of what is being learnt. Even though this definition is currently very controversial (e.g., Reber, 1994; Shanks & StJohn, 1994), there is now a large body of evidence suggesting that improvements in performance at a given task are not systematically accompanied by similar improvements in subjects's ability to express or use the acquired knowledge in an explicit way. For instance, Artificial Grammar Learning studies have shown that subjects can classify strings of letters as grammatical or not after practice at memorizing similar strings, and without being able to report on the rules that define grammaticality (e.g., see Dienes & Berry, 1994, for a review). Sequence learning studies, on which this paper will focus, have demonstrated that subjects can become sensitive to the regularities contained in sequences of stimuli presented in a choice reaction setting despite remaining unable to report on the sequence or to perform well in other direct tests such as generation, where subjects are asked to predict the next stimulus instead of reacting to the current one (e.g., Cleeremans, 1993a; Nissen & Bullemer, 1987).

Implicit learning is assumed to have a number of features that distinguish it from explicit learning. However, because the existence and nature of implicit learning is controversial, there is currently no agreement in the field about which features have been empirically established. For instance, some authors claim that implicit learning is an unconscious process that can result in abstract knowledge (e.g., Reber, 1994). For others, however, implicit learning is essentially explicit exemplar-based learning (e.g., Shanks & StJohn, 1994). These issues have been extensively explored empirically and I believe that it is fair to say that they remain largely unsolved at this point.

In this paper I would like to focus on a assumption that often underpins the others but that has seldom been addressed directly, that is, that implicit learning is an independent and automatic process. Three positions about this issue have been expressed in the implicit learning literature.

First, some authors (e.g., Perruchet & Amorim, 1992) argue that performance in implicit learning tasks does not necessarily reflect the operation of an independent implicit learning system. Rather, performance would be mostly based on explicit processing, but the resulting knowledge is fragmented enough that verbal reports probing for general information are unlikely to reveal the extent of subjects' knowledge.

Other authors (e.g., Knowlton, Ramus & Squire, 1992) assume that implicit and explicit learning are supported by different memory systems, and that these systems are completely independent from each other. Implicit and explicit learning would thus proceed in parallel, and without interacting. They produce different kinds of knowledge, and are most likely to operate efficiently in contrasted settings.

Finally, there may be an intermediate position where one assumes that implicit and explicit processing indeed rely on distinct memory systems, but in which some interactions between the two systems are allowed, and in which some processing resources are shared (e.g., Cleeremans, 1993b).

Typically, these issues have been approached by placing subjects in dual task settings. For instance, Keele and his collaborators (e.g., Curran & Keele, 1992) used sequential reaction time (SRT) tasks coupled with a secondary tone-counting task. The rationale of these experiments was to determine whether learning of the sequential structure of the stimulus material can still occur despite attentional resources being recruited by the secondary task. In other experiments Curran & Keele also manipulated subject's explicit knowledge by letting them study the sequence beforehand.

[3]The author is a Research Associate of the National Fund for Scientific Research (Belgium).

In general, the results of these and other studies have shown that the availability of explicit knowledge results in better performance under single task conditions, but that the presence of a secondary task results in all subjects performing at the same level regardless of whether or no they possess explicit knowledge.

This kind of result has often been interpreted as yielding support to the notions (1) that sequence learning involves both explicit and implicit learning, and (2) that implicit learning relies on mechanisms that are independent from short-term memory and from the availability of attentional resources, both of which are crucially important for explicit learning to occur.

This methodology has a number of problems, however. First, it is difficult to assess how much attentional capacity the secondary task requires. For instance, the difficulty of keeping track of how many tones of a particular kind have been presented so far may vary with the number of tones presented. This problem makes it hard to draw strong inferences about the relative independence of the processes responsible for learning the sequential structure.

Second, it is hard to determine how much explicit knowledge subjects possess and actually use during the task.

In this paper, inspired by work on overshadowing in conditioning experiments with animals (e.g., Matzel, Schachtman & Miller, 1985), I report on a different way to address this issue. Instead of placing subjects in a dual-task setting, I placed them in a dual-stimulus setting where on each trial, two sources of information are available to compute the response: the sequential context set by previous elements of the sequence on the one hand, and a cue indicating where the next stimulus will appear on the other hand. Hence, on each trial subjects can rely on either or both the temporal context and the information conveyed by the cue to prepare for the next event. Do subjects learn about both dimensions or is learning of one dimension blocked by learning of the other? If implicit learning truly is an automatic process that relies on independent processing and memory systems, then one should expect to obtain such learning even in conditions where more reliable and fully explicit sources of information about the relevant material are available.

In the following, I first describe experimental work that implements the design outlined above. Next, I report on preliminary simulation work that illustrates how one can explore these issues within the connectionist framework.

## Experimental Design

### Method

The experiment consisted of three conditions. Each consisted of 10 training sessions during which subjects were exposed to a six-choice SRT task. Each session consisted of 20 blocks of 150 trials each, for a total of 30,000 trials. After training all subjects were exposed to 3 blocks of a generation task in which they were asked to predict the location at which the next stimulus would appear. On each trial of the SRT task, a stimulus could appear at one of six positions arranged horizontally on a computer screen, and subjects were to press as fast and as accurately as possible on the key corresponding to the current location of the stimulus. Subjects were kept unaware that the material was structured sequentially. The sequential structure of the material was manipulated by generating the sequence based on a noisy finite-state grammar, as described below. In cued blocks, a cue consisting of a cross under one of the six stimulus positions appeared concurrently with the stimulus. This cue could either be valid or invalid. If valid, it indicated the location at which the next stimulus in the sequence would appear.

During generation, the same stimulus material was presented, but subjects were instructed to try to predict the next stimulus instead of merely reacting to the current one. No explicit feedback was provided during generation performance to minimize within–generation learning.

In the Low Validity (LV) condition, each session consisted of 2700 cued trials followed by 300 neutral trials. Cue validity was set at 20%. In the High Validity (HV) condition, the same design was used but cue validity was considerably higher (80%). The third condition (100% validity, or HV100) followed a somewhat different design: Each of the first 9 sessions consisted of 3000 trials for which cue validity was 100%, and the final session consisted exclusively of 3000 neutral trials.

### Subjects

Six subjects participated in each of the three experiments. Subjects were paid about $65 for participating in the experiment, and could earn an additional bonus of $34 to $62 based on performance in the SRT task.

### Apparatus and Display

The experiment was run on Macintosh computers. The display consisted of six dots arranged in a horizontal line on the computer's screen and separated by intervals of 3 cm. Each screen position corresponded to a key on the computer's keyboard. The spatial configuration of the keys was fully compatible with the screen positions. The stimulus was a small black circle 0.35 cm high that appeared on a white screen background, centered 1 cm below one of the six dots. The cue was a small cross (X) appearing at the same locations as the stimuli. The RSI was 120 msec.

### Stimulus and cue generation

Stimuli were generated based on the noisy finite-state grammar illustrated in Figure 1, with a small proportion of random stimuli (20%) interspersed with those derived from the grammar. Learning is assessed by comparing performance on stimuli that follow the rules of the grammar versus random stimuli. A total of 30,000 trials were presented to each subject. On each trial, stimulus generation proceeded in three phases. First, an arc coming out of the current node was randomly selected, and its label recorded. The current node was set to be #0 on the first trial of any block, and was updated on each trial to be the node pointed to by the selected arc. Second, there was a 20% chance of substituting a randomly selected label to the recorded one (identity substitutions, as well as any substitution that would result in a stimulus being repeated or legal at the current

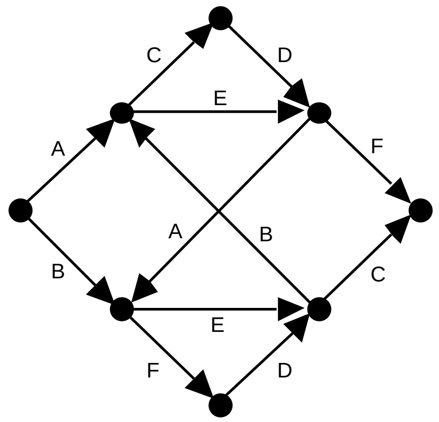


**Figure 1:** Finite-state grammar (from Jiménez & Cleeremans, 1994) used to generate the stimulus material. Note that the first and last nodes are one and the same.

node, were not allowed). Third, the label was used to determine the screen position at which the stimulus would appear by following a 6 x 6 Latin square design, so that each label corresponded to each screen position for exactly one of the six subjects in each condition. Note that each label appears twice in the grammar and may be followed by different successors on different occurrences. Maximally reducing the uncertainty associated with each label requires encoding up to three elements of temporal context. Cue generation proceeded independently. On each trial, a cue corresponding to the next stimulus was generated. This valid cue could be presented on all trials of a given session (first 9 sessions of the HV100 condition), suppressed entirely (neutral blocks or sessions) or be replaced by an invalid cue in either 20% (HV condition) or 80% (LV condition) of the trials. Substitution consisted of selecting a random location for the cue to appear at, with the constraints that this location could be neither the location of the current stimulus location nor that of the next one.

To summarize, this generation procedure results in six types of trials defined by crossing the Grammaticality (Grammatical or Non-Grammatical) and Cue Validity (Valid, Invalid, or Neutral) factors. A particular trial was thus categorized as “valid” if the location at which the stimulus had appeared on that trial had indeed been validly primed by the cue that had appeared on the previous trial. Similarly, a given trial was categorized as “grammatical” if the stimulus that had appeared on that trial was consistent with the generation rules expressed by the finite-state grammar.

Finally, note that the sequence generation procedure makes it impossible for any stimulus to be involved in a direct repetition of itself. This guarantees that RT effects are not contaminated by short-term priming effects, which have large facilitatory effects on performance that are completely independent from the factors of interest in this research.

## Results and Discussion

Figure 2 represents average RTs over the 10 sessions of training and for each trial type, in the LV (left panel), HV (middle panel) and HV100 (right panel; note the scale difference) conditions. Consider first the data for the HV condition (middle panel). It is obvious that cue validity has a large effect on performance, as RTs elicited by valid trials are considerably faster than those elicited by both neutral and invalid trials. This pattern of results indicates that subjects are indeed using the cue to anticipate the location at which the next event will appear and to specifically prepare their response accordingly. Despite the massive impact of cue validity, small effects of grammaticality are also present at all levels of cue validity, and seem to have approximately the same magnitude in each case. These impressions were confirmed by an ANOVA with Practice (10 levels) X Validity (Valid, Invalid or Neutral) X Grammaticality (Grammatical vs. Non-Grammatical) as factors and RT as dependent variable. The analysis yielded significant main effects of Practice [$F(9, 45) = 27.389$, $p < .001$, MSe = 64494.037], of Cue Validity [$F(2, 10) = 162.385$, $p < .001$, MSe = 1849184.178], and of Grammaticality [$F(1, 5) = 39.685$, $p < .01$, MSe = 29322.225], as well as a significant interaction between Practice and Cue Validity [$F(18, 90) = 8.680$, $p < .001$, MSe = 7443.181]. There was no significant interaction between Cue Validity and Grammaticality ($p > .05$).

Overall, these results suggest that sensitivity to sequential structure was not blocked by the presence of the cue, despite the facts that subjects (1) demonstrably use the cue, and (2) that the cue conveys considerably more reliable information about the next event than the sequential structure does.

Unsurprisingly, cue validity has a much smaller impact on performance in the LV condition (left panel): Valid, invalid and neutral trials elicit similar RTs over training (with the exception of neutral trials in the first session, see below). Grammatical trials at all levels of cue validity, however, elicit faster RTs than non-grammatical trials, just as in the HV condition. Thus, subjects do not appear to use the unreliable cue in the LV condition, relying instead on the sequential structure to optimize their performance at the task. This analysis was again confirmed by an ANOVA, which produced significant main effects of Practice [$F(9, 45) = 54.328$, $p < .001$, MSe = 157235.027] and of Grammaticality [$F(1, 5) = 17.987$, $p < .01$, MSe = 123839.803]. Surprisingly, Cue Validity was also significant [$F(2, 10) = 26.799$, $p < .001$, MSe = 14187.900] and interacted with Practice [$F(18, 90) = 5.098$, $p < .001$, MSe = 3479.17]. Closer examination of the figure reveals that these effects are in fact artifactual. Indeed, the neutral trials presented during the first session elicit much faster RTs than either valid or invalid trials. However, this is merely a result of the fact that these trials were presented at the end of the first session. Hence they benefit from previous unspecific training on the other trials during the first session. This artifact is absent from the subsequent sessions, and analyses that exclude the first session produce non significant effects of Cue Validity. Finally, in contrast with the HV data, the ANOVA also revealed a significant interaction between Grammaticality and Practice [$F(9, 45) = 3.311$, $p < .01$, MSe = 601.642].

A further ANOVA on the data from both conditions revealed significant main effects of Practice [$F(9, 90) = 80.780$, $p < .001$, MSe = 210466.871], of Cue Validity [F(2,

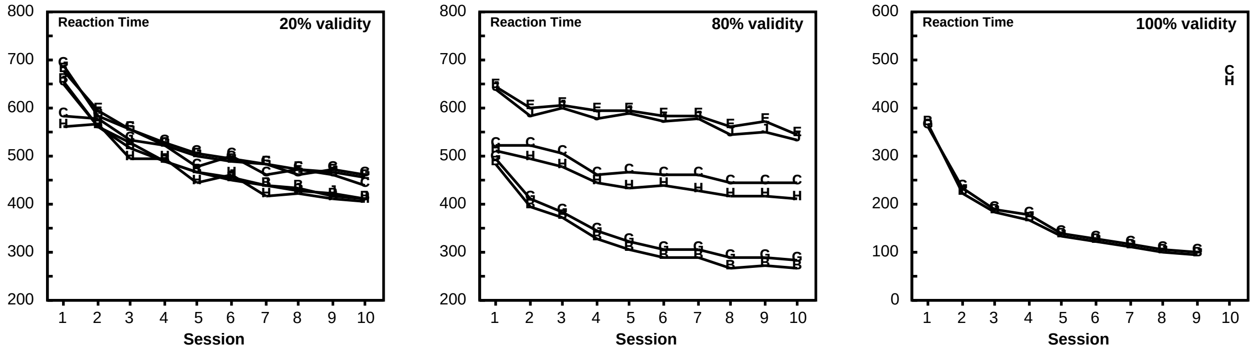


**Figure 2:** Reaction times for the Low Validity (left panel), High Validity (middle panel) and 100% Validity (right panel) conditions. RTs are represented separately for valid (squares), invalid (circles), and neutral (triangles) cues, as well as for grammatical trials (filled symbols) and non-grammatical trials (open symbols).

20) =155.355, p < .001, MSe = 925690.239], and of Grammaticality [F(1, 10) = 35.898, p < .001, MSe = 136840.939]. Condition failed to reach significance. Cue validity interacted significantly with Condition [F(2, 20) = 157.367, p < .001, MSe = 937681.239], but not Grammaticality (p = .0654).

To summarize these data, subjects appear to remain sensitive to the sequential structure even when a far more reliable source of information is available to anticipate the next event. Recall that in the HV condition, the cue will reliably predict the next event in 80% of the cases. By contrast, a simple examination of the FSA illustrated in Figure 2 shows that even full knowledge of a *deterministic* version of the grammar would only allow for about 50% of the trials to be correctly anticipated, as most nodes have two equiprobable outgoing arcs.

The fact that neutral trials were interspersed throughout training makes strong inferences difficult, however. It remains possible that these trials provided enough training for subjects to learning about the sequential structure of the material.

To test this hypothesis, a further condition was run in which subjects were exposed exclusively to valid cued trials throughout the first 9 sessions, and subsequently transferred to a final session consisting exclusively of neutral trials. The corresponding data are presented in the rightmost panel of Figure 2. One can see that even though grammatical and ungrammatical trials fail to elicit different RTs throughout training, a significant difference of 23 msec [One-tailed t(4) = 2.79, p < 0.1] reappears in the last session.

To determine whether this difference could be attributed to learning within the last session, I conducted this analysis again but restricted it to the first 300 trials of this session. The difference between grammatical and ungrammatical trials now averaged 30.2 msec and was significant [One-tailed t(4) = -4.371, p < 0.1]. This suggests that subjects did acquire knowledge about the sequential structure of the material during training but were unable to express it because of the presence of the cue.

Finally, Figure 3 represents reaction times to grammatical and ungrammatical neutral trials during the last sessions of all three conditions. The figure shows that these differences tend to be very similar in all three conditions. An ANOVA conducted on these data confirmed this impression, with a significant main effect of Grammaticality [F(1, 15) = 45.844, p < .001, MSe = 2976.250] and no interaction between Grammaticality and Condition (p = 0.57). Overall then, subjects appear to learn about the sequential structure of the material regardless of the validity of the cue.

Space limitations prevent a full treatment of the generation task data, but subjects were consistently unable to better predict grammatical elements over ungrammatical elements. This indicates that whatever knowledge was acquired over training with the RT task was of limited use in helping subjects produce explicit prediction responses.

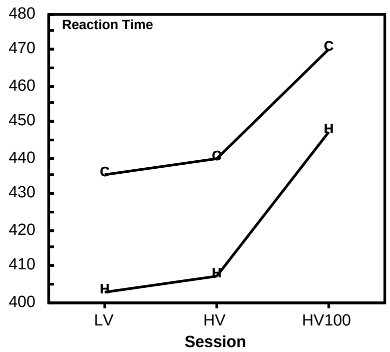


**Figure 3:** Reaction times for grammatical (filled symbols) and ungrammatical (open symbols) neutral trials presented during the last session of each the three conditions (LV: Low validity; HV: High Validity; HV100: 100% validity).

## Simulation Results

What kind of mechanism may account for these data? A natural starting point is the Simple Recurrent Network (SRN) model first proposed by Elman (1990), and shown in Figure 4 (inside the frame). The task of this back-propagation network is to predict the next element of a sequence based on the current element and on a representation of the temporal context that the network has elaborated itself. Over training, the network's responses come to approximate the optimal conditional probabilities associated with each successor to the current context, and can thus be interpreted as representing preparation for the next event. Previous work (see Cleeremans & McClelland, 1991; Cleeremans, 1993a, for detailed analysis of both processing in such networks and

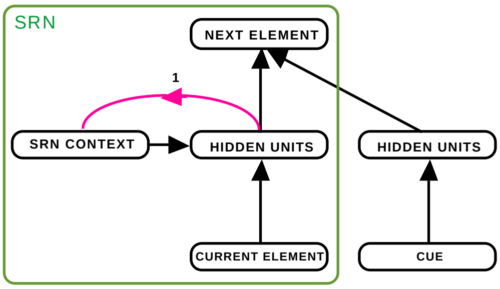


**Figure 4:** The simple recurrent network (framed) augmented with an additional pathway to process cue information.

correspondence with human data) has shown that the SRN is able to account for about 80% of the variance in SRT data.

To model performance in the experiments described above it is necessary to augment the SRN architecture with mechanisms that enable it to process the information conveyed by the cue. There are several different ways of doing this according to which assumptions one has about the way in which learning about the cue and learning about the sequential structure of the material interact. First, one may assume that processing of the cue is fully independent from processing of the sequential structure. Thus, learning proceeds independently in each processing pathway, just as if two separate networks were trained independently, and information conveyed by each pathway is only combined at response time.

Second, one may assume that learning of one dimension interacts with learning of the other dimension. This is the case in backpropagation architectures where both pathways feed into a single output module. Indeed, in such arhitectures, the pathway that transmits more information will tend to develop larger connection weights and exert an increasingly larger influence on performance, at the expense of the other pathway.

The network represented in Figure 4 is an instance of this latter class of models: The SRN is simply augmented with an additional processing pathway consisting of input units to represent the cue. These units feed into a set of hidden units which are in turn connected with the output units.

To assess how well this kind of network was able to account for SRT performance in this experiment, I conducted simulations in which the model was trained on the same material as human subjects and for the same number of trials, with the parameters used by Cleeremans and McClelland (1991). The network used local representations on both the input and output pools (i.e., each unit corresponded to one of the 6 stimuli or cues). To account for short-term priming effects, the network used dual connection weights and running average activations on the output units, as described in Cleeremans and McClelland (1991).

The network was trained to predict each element of a continuous sequence of stimuli generated in exactly the same conditions as for human subjects. On each step, both a label and a cue were generated as described before and presented to the network by setting the activation of the corresponding input units to 1.0. Activation was then allowed to spread to the other units of the network, and the error between its response and the actual successor of the current stimulus was then used to modify the weights.

During training, the running average activation of each output unit was recorded on every trial and transformed into Luce ratios (Luce, 1963) to normalize the responses. For the purpose of comparing the model's and the subject's responses, I assumed (1) that the normalized running average activations of the output units represent response tendencies, and (2) that there is a linear reduction in RT proportional to the relative strength of the unit corresponding to the correct response. The network's responses were subtracted from 1.0 to make increases in response strength compatible with reduction in RT.

The resulting data are shown in Figures 5 and 6. One can see that the model's performance approximates human performance quite well, at least qualitatively. Indeed, just as human subjects, the model appears to be sensitive to the sequential structure of the material at all levels of cue validity in all three conditions. The relative size of the differences between performance on neutral trials during the last session of each of the three conditions is also well preserved in the simulations (Figure 6).

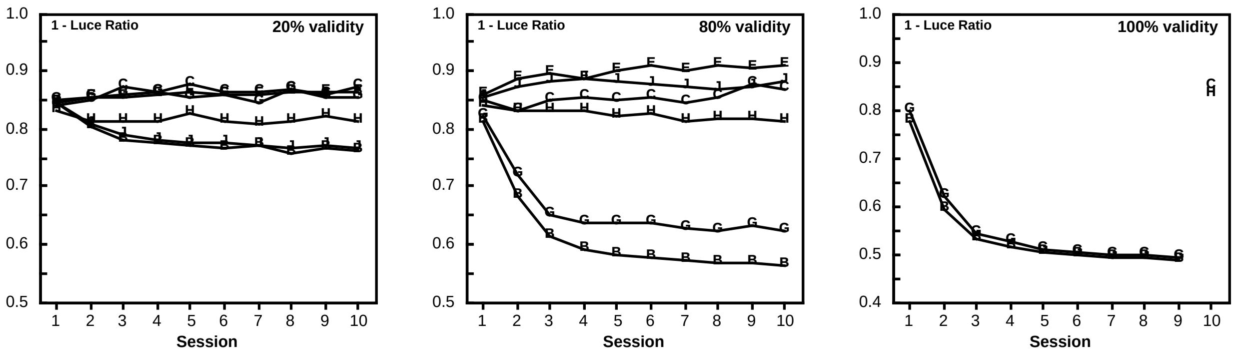


**Figure 5:** Simulated SRN responses (see text for details) for the Low Validity (left panel), High Validity (middle panel) and 100% Validity (right panel) conditions. Responses are represented separately for valid (squares), invalid (circles), and neutral (triangles) cues, as well as for grammatical trials (filled symbols) and non-grammatical trials (open symbols).

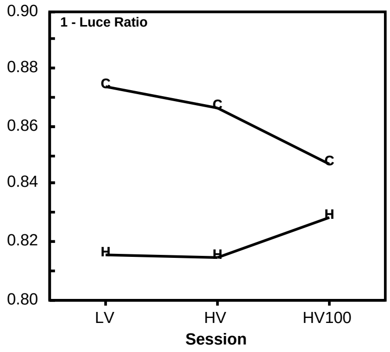


**Figure 6:** Simulated SRN responses (see text for details) for grammatical (filled symbols) and ungrammatical (open symbols) neutral trials presented during the last session of each the three conditions (LV: Low validity; HV: High Validity; HV100: 100% validity).

Note however that there are also discrepancies. In particular, the model fails to account for unspecific practice effects, that is, changes in reaction times that do not result specifically from the presence of sequential structure. This is a flaw shared by previous versions of the SRN model, but one that is not crucial to the arguments develoved in this paper.

Work in progress is aimed at contrasting these results with those produced by architectures in which the two processing pathways are trained completely independently.

## Conclusion

In this paper I presented three experiments aimed at exploring to what extent implicit learning of sequential structure in RT tasks proceeds independently of the availability of explicit knowledge about the stimulus material. By contrast to standard dual-task procedures that have been used in the past to explore this issue, I used a dual-stimulus setting where on each trial, subjects were exposed to both the current stimulus and to a cue of varying validity that indicated where the next stimulus would appear. The results indicated that even in conditions where the cue was a much better source of information about the next event, subjects still seemed to be sensitive to the sequential structure of the material. Hence acquisition of sequential structure can proceed even in conditions where vastly superior and fully explicit sources of information about relevant task information are available.

However, simulation work using a model based on the Simple Recurrent Network indicated that these results also obtain in architectures where the two processing pathways (sequential structure Þ next event, and cue Þ next event) are not fully independent. This suggests that preserved learning along one dimension does not necessarily entail that the underlying structures are themselves fully independent.

Further empirical research and modeling work is needed to increase our understanding of the relationship between performance and underlying processing modules, but the empirical data clearly suggest that implicit learning of sequential structure is a resilient process that is little sensitive to the availability of other, more reliable, and fully explicit sources of information about the stimulus material.

## Acknowledgments

The author is a Research Associate of the National Fund for Scientific Research. I thank Marie-Catherine Navarre and Arnaud Destrebecqz for help in data collection and an anonymous reviewer for valuable comments.

## References

Cleeremans, A. (1993a). *Mechanisms of implicit learning: connectionist models of sequence processing*. Cambridge, MA: MIT Press.

Cleeremans, A. (1993b). Attention and awareness in sequence learning. In *Proceedings of the Fiftheenth Annual Conference of the Cognitive Science Society* (pp. 330–335). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cleeremans, A. & McClelland, J.L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120, 235–253.

Curran, T. & Keele, S.W. (1992). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 189–202.

Dienes, Z. & Berry, D. (1994). *Implicit learning: Theoretical and empirical issues*. Hove, England: Lawrence Erlbaum Associates.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.

Jiménez, L., & Cleeremans, A. Direct and indirect measures of sequence learning. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 445-450). Hillsdale, NJ: Lawrence Erlbaum Associates.

Knowlton, B.J., Ramus, S.J. & Squire, L. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science*, 3, 172–177.

Luce, R.D. (1963). Detection and recognition. In R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (vol. 1). New York: Wiley.

Matzel, L.D., Schachtman, T.D. & Miller, R.R. (1985). Recovery of an overshadowed association by extinction of the overshadowed stimulus. *Learning and Motivation*, 16, 398-412.

Nissen, M.J. & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32.

Perruchet, P. & Amorim, P.A. (1992). Conscious knowledge and changes in performance in sequence learning: evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 785–800.

Shanks, D.R., & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367–395..

Reber, A.S. (1994). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford, England: Oxford University Press.