# Implicit Learning and Consciousness

## An empirical, philosophical and computational consensus in the making

edited by

Robert M. French
*Department of Psychology*
*University of Liège*

Axel Cleeremans
*Cognitive Science Research Unit*
*Université Libre de Bruxelles*

(2002)

PSYCHOLOGY PRESS • Taylor & Francis Group •

CHAPTER TWO

# The Self-organising Consciousness: A framework for implicit learning

**Pierre Perruchet and Annie Vinter**
*Universite de Bourgogne, France*

The prevalent models of implicit learning have difficulty in accounting for the critical importance of the attention paid to the study material during the familiarisation phase of a learning session. In this chapter, we show that a close examination of the on-line content of successive attentional focuses, which form subjects' phenomenal experience, suggests a new interpretation of implicit learning. This interpretation is based on the fact that conscious contents are self-organising. We first present the notion of Self-organising Consciousness in the context of the discovery of words in artificial languages, and then generalise it to the formation of other forms of conscious representations. We then discuss how the concept of Self-organising Consciousness allows us to think in a non-standard way about the processes occurring in implicit learning situations, and even in situations involving some form of abstraction with regard to the surface features of the material. Finally, we discuss the meaning and the validity of introducing consciousness into a causal, computationally implementable framework.

## ATTENTION IS A CONDITION FOR LEARNING

Most overviews of the literature (e.g. Cleeremans, Destrebecq, & Boyer, 1998) distinguish between three main situations in which implicit learning has been studied. In all three cases, subjects have to deal with a situation governed by complex, arbitrary rules, without being prompted for an explicit analysis of the task.

In artificial grammar learning (e.g. Reber, 1976), the material is usually composed of a set of consonants, the nature and the ordering of which are defined by a finite-state grammar. Subjects have to learn by rote the letter strings generated by this grammar. After the training session, they are asked to judge the grammaticality of new sequences, half of which obey the rules of the grammar and half of which violate them.

In the dynamic system control tasks (e.g. Berry & Broadbent, 1988), subjects have to control a dynamic system simulated on a computer, such as a sugar production factory or a city transport system. They are required to reach and maintain specified target values of an output variable by varying one or two input variables. The output and input variables are linked by a linear equation including the current state of the system as one term.

Finally, in sequential reaction time tasks (e.g. Nissen & Bullemer, 1987), subjects have to respond to targets located at several places on a computer screen. In most cases, the underlying rule is simply that one and the same sequence is continuously repeated. Performance is assessed through choice reaction times to these targets.

There are many variants of these prototypical situations, and other situations have received increasing attention in recent years (for instance the implicit learning of invariant characteristics in the McGeorge & Burton [1990] paradigm). Whatever the paradigm, however, the same two results have emerged. Firstly, subjects are shown to perform above chance levels in such situations; secondly, most aspects of this behavioural adaptation to the structural features of the situation are not due to the intentional exploitation of subjects' explicit knowledge about these features (see Chapter 1 for a more detailed presentation of this literature).

A question of major interest is whether performance improvement depends on the amount of attention paid to the study material during the familiarisation phase. This question has not received the same amount of interest in the different subareas of the field, but, in each case, the same methodology has been applied. The principle consists in adding a concurrent secondary task during the training session, then observing whether performance improvement is equivalent to that observed in a single-task, standard procedure. Regarding artificial grammars, Reber (e.g. 1993) has acknowledged that attention to the study material is necessary for learning to occur, although the point has rarely been addressed in empirical investigations. In support of this claim, Dienes, Broadbent, and Berry (1991) have shown that the accuracy of grammaticality judgements was lowered when subjects had to perform a concurrent random number generation task during the familiarisation phase.



The studies carried out on dynamic system control tasks provide a more contrasted picture. The theory surrounding the early studies on these tasks posited a distinction between two forms of learning: Selective (i.e. with attention) and unselective (i.e. without attention; e.g. Berry & Broadbent, 1988). Unselective learning was assumed to occur when the situation was too complex to be solved by attention-based mechanisms. In support of this conception, Hayes and Broadbent (1988) reported that adding a secondary task of letter or digit generation interfered with learning a simple system control task, but facilitated learning in a more complex version of the same task. However, these results have been criticised for the lack of power of the statistical tests (Dienes et al., 1991). Moreover, both Dulany and Wilson (unpublished data) and Green and Shanks (1993) failed to replicate these results (despite extensive attempts to do so), and observed that, as a rule, the secondary task impaired performance irrespective of the complexity of the task. To the best of our knowledge, the notion of unselective learning, as initially discussed in the studies conducted by Broadbent and colleagues, is no longer advocated.

The idea of two forms of learning, differing according to whether attention is required or not, has also been proposed in the context of repeated sequence learning, although the proposal is diametrically opposed to the position adopted by Broadbent and co-workers. Indeed, the hypothesis was that attention is required for learning complex sequences, while non-attentional learning is efficient for the simplest forms of sequential dependencies. The secondary task typically used in this context is a tone-counting task, in which a high- or low-pitched tone sound is emitted after each trial and subjects are required to keep a running count of one of them (e.g. high-pitched tones). According to Cohen, Ivry, and Keele (1990), introducing this secondary task impaired learning of complex sequences, while failing to affect the learning of sequences incorporating only first-order contingencies.

However, again, this view no longer appears to be tenable. Observing performance improvement under dual-task conditions does not imply the existence of a non-attentional form of learning, because the secondary task might not deplete the attentional resources completely. As claimed by Stadler (1995, p. 683) "Even when implicit serial learning is observed in conjunction with the tone-counting task, . . ., it cannot be said that learning occurred without attentional capacity—the participants certainly devoted attention to the serial reaction time task". Evidence favouring attentional involvement stems notably from the fact that, as a rule, normal subjects have acquired explicit knowledge of the sequence after the training session, even when training was performed under dual-task conditions (e.g. Perruchet & Amorim, 1992; Shanks & Johnstone, 1999). The logic here is that because explicit memories require attentional processes, recall or recognition of a stimulus can be used as a measure of prior stimulus attendedness (e.g. Schmidt & Dark, 1998, p. 228). Closing their recent survey of the role of attention in implicit sequence learning, Hsiao and Reber (1998, p. 487) concluded: "We view sequence learning as occurring in the background of the residual attention after the cost of the tone-counting task and the key-pressing task. If there is still sufficient attention available to the encoding of the sequence, learning will be successful; otherwise, failure will result"; for other approaches that emphasise the role of attention, see Frensch, Buchner, and Lin (1994) and Jiménez and Mendez (1999). Note that this conclusion does not imply that performance is systematically impaired in dual-task conditions. Indeed, in a few cases, performance has been found to be insensitive to capacity load (Jiménez & Mendez, 1999; Stadler, 1995), a result that can be easily accounted for by some kind of floor or ceiling effect.

Thus, recent studies strongly challenge the claims that two forms of learning can be distinguished, with a non-attentional form emerging when the situation is very complex (e.g. Berry & Broadbent, 1988) or very simple (Cohen et al., 1990). Most probably, improved performance in any implicit learning situation implies at least minimal attentional involvement. This conclusion comes as no surprise, because the major role played by selective attention in acquisition processes is an old and robust empirical finding. This role has been identified, for instance, in the literature on animal conditioning (Mackintosh, 1975), automatisms (Fisk & Schneider, 1994), perception (Rock & Gutman, 1981), and memory (Roediger, 1990). This contention holds even for the so-called implicit memory phenomena, in which performance does not involve the recollection of the initial episodes. There is now overwhelming evidence that attention to the material at the time of encoding is a necessary condition for the observation of improved performance in subsequent implicit memory tests, such as word completion and perceptual identification tasks (Crabb & Dark, 1999), reading tasks (MacDonald & MacLeod, 1998), or object decision tasks (Ganor-Stern, Seamon, & Carrasco, 1998).

## FROM ATTENTIONAL RESOURCES TO PHENOMENAL CONSCIOUSNESS

Attention is generally construed in terms of mental resources or capacity. The usual way of manipulating attention in implicit learning settings (namely, the dual task paradigm) and the aim of examining whether performance is impaired when attention is shared between different sources of information, are in keeping with this energetic, quantitative view of attention. Within this perspective, the need for a certain amount of mental resources for learning in the complex learning settings appears to be anything but surprising.



However, a closer look at this issue reveals a striking paradox. The general view of most contributors to the implicit learning literature is that implicit learning relies on powerful unconscious mechanisms, the main property of which is, precisely, their freedom from capacity limitations. Thus the idea that the involvement of attention is necessary for the recruitment of mechanisms that are thought to be free of attentional limitations is somewhat awkward.

This paradox appears still more pronounced if, instead of considering attention as a pool of processing capacities, we consider its content, which is generally identified with subjects' phenomenal consciousness (e.g. Cowan, 1995; Mandler, 1975; Miller, 1962; Posner & Boies, 1971).[1] Indeed, it turns out that, whatever this content might be, it certainly does not match the cognitive operations that most researchers hypothesise. Even a rough post-experimental interview is sufficient to make it clear that attention is not devoted to abstracting the rules that the advocates of abstractionist positions (e.g. Marcus, Vijayan, Rao, & Vishton, 1999; Reber, 1993) consider essential. Likewise, attention is certainly not devoted to the statistical analysis of the raw data that other authors claim to be the source of performance improvement in implicit learning settings (e.g. Cleeremans, 1993). Attention, presumably, is centred on the on-line processing of the sensory data. Thus, considering the object of attention actually makes the necessity of attention in implicit learning still more puzzling than considering attention only as an unspecified pool of capacity. Indeed, it looks quite mysterious why the initial coding of the sensory data should need attention, while all the subsequent operations performed on these data could be run without any attentional engagement and conscious counterparts.

To summarise: (1) empirical data provide evidence of a close link between improved performance in implicit learning settings and the amount of attention devoted to the material during the familiarisation phase of the task;

[1] Some authors (e.g. Velmans, 1999) have argued that phenomenal consciousness and attention ought to be distinguished, because attention is selective whereas "consciousness incorporates both a central focus, and a rich polymodal periphery", to borrow the expression used by O'Brien and Opie (1999). This argument amounts to defining attention as the conceptually driven mechanisms that are directed towards a specific source of information in response to task instructions. This view defines what Schmidt and Dark (1998) call the intention-equals-attention view, according to which participants' intention to attend exclusively to a target is sufficient to restrict attentional processing to this target. However, the fact that the instructions ask participants to pay attention to a target does not prevent them from making quick attentional shifts toward non-attended information. Therefore, unless one endorses a highly restrictive definition of attended information as the informational content on which subjects are asked to focus, we see no reason to dissociate attention and consciousness on the basis of their relative selectivity.

and (2) prevalent theories of implicit learning fail to incorporate this finding. Indeed, the dependence of implicit learning on limited attentional capacity is contradictory to the postulate (central to such theories) of a powerful unconscious processor, and this contradiction is exacerbated if one considers the mismatch between the content of attention such as it is revealed through verbal reports and the hypothesised mental operations. Overall, the hypothesised machinery would appear to be particularly ill adapted. If evolution has endowed human minds with the powerful and unlimited analytical capabilities that most theoreticians of implicit learning attribute to the unconscious mind, it appears especially inefficient that the expression of these capabilities should be dependent on an attentional bottleneck acting on the presumably simpler perceptual processing of the information.

Before taxing evolution for having missed its shot, however, we should perhaps examine whether the error might lie with the theories of implicit learning. The present chapter is intended to show that if learning needs attention, then this is simply because learning is a natural by-product of the attentional, on-line perceptual processing of the incoming information. The powerful processors whose existence is postulated by researchers, whether they are thought to be in the service of rule abstraction or statistical analysis, are devoid of any object. The subsequent sections are aimed at explaining why the close scrutiny of conscious contents in a dynamic perspective provides us with a new explanation for performance improvement in implicit learning. This is because this conscious content, as we will shortly show, has the astonishing property of being capable of self-organisation.[2]

## THE SELF-ORGANISATION OF PERCEPTUAL EXPERIENCE IN WORD DISCOVERY

We introduce the notion of Self-organising Consciousness (SOC) by starting from the discovery of the words that form oral language as an

[2] At this point, it is crucial to avoid a deep misunderstanding of our proposal. In the implicit learning area, the term "consciousness" is regularly associated with the explicit knowledge that subjects might have gained about the study material during the study phase. As a rule, studies in which the measures of explicit knowledge fulfill both Shanks and St John's (1994) information and sensitivity criteria show that performance improvement is accompanied by the explicit knowledge of the relevant aspects of the procedure. In this context, our conception could be expected to be a variant of the view that subjects exploit their explicit knowledge about the structure of the situation with the powerful analytical tools of conscious thought in order to anticipate the next event or to select the right response in each situation. Our conception is, in fact, radically different. In the present chapter, the term "consciousness" designates the on-line content of the attentional focus, and not the explicit knowledge that subjects might have developed about the material, such as it might be revealed in post-experimental tests.



example of implicit learning. The reason we do not take one of the three major paradigms discussed above as our starting point is convenience of presentation, and it is worth stressing from the outset that the argument below is also relevant to these classical paradigms, in a way that will be made clear in a subsequent section.

Language acquisition initially proceeds from auditory input, and linguistic utterances usually consist of sentences linking several words without clear physical boundaries. The question thus arises: How do infants become able to segment a continuous speech stream into words? Recent psycholinguistic research has identified a number of prosodic and phonological cues that could potentially help infants, but they provide only probabilistic information. The importance of prosodic and phonological cues in word discovery is further questioned by recent experimental studies showing that these cues are not necessary. For instance, Saffran, Newport, and Aslin (1996a) used an artificial language consisting of six trisyllabic words, such as *babupu* and *bupada*. The words were read by a speech synthesiser in random order in immediate succession, without pauses or any other prosodic cues. Thus the participants heard a continuous series of syllables without any word boundary cues. In the following phase, they were asked to perform a forced choice test in which they had to indicate which of two items sounded more like a word from the artificial language. One of the items was a word from the artificial language, whereas the other was a new combination of three syllables belonging to the language. Participants performed significantly better than would be expected by chance. This and other studies (Saffran, Aslin, & Newport, 1996b; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) offer impressive support for the hypothesis that people are able to learn the words forming a continuous speech stream without any prosodic or phonological cues for word boundaries. Our aim in this section is to show that word extraction can be explained by the action of elementary, associative-like processes acting on the initial conscious percepts, the result of which is to modify the conscious experience we have of the linguistic input.

What is the phenomenal experience of the listener of a new language such as the one used in the Saffran et al. experiments, at the beginning and end of training respectively? When people are confronted with material consisting of a succession of elements, each of them matching some of their processing primitives, they segment this material into small and disjunctive parts comprising a small number of primitives. As adults, we have direct evidence of the phenomenon. For instance, when asked to read nonsense consonant strings, we read the material not on a regular, rhythmic, letter-by-letter basis, but rather by chunking a few letters together. The same phenomenon presumably occurs when a listener is faced with an unknown spoken language, with the syllables or other phonological units forming the subjective processing primitives instead of the letters. Chunking, we contend, is a ubiquitous phenomenon, due to the intrinsic constraints of attentional processing, with each chunk corresponding to one attentional focus. Importantly, however, the listener's initial conscious experience consists of a succession of chunks that have only a weak probability of matching the words of the language.

After extensive exposure to the language, the listener's phenomenal experience is presumably the experience each of us has of our mother tongue, that is the experience of perceiving a sequence of words. Our proposal is that the final phenomenal experience of perceiving words emerges through the progressive transformation of the primitives guiding the initial perception of the language, and that this transformation is due to the self-organising property of the content of phenomenal experience. The basic principle is fairly simple. The primitives forming a chunk, that is, those that are perceived within one attentional focus as a consequence of their experienced temporal proximity, tend to pool together and form a new primitive for the system. As a consequence, they can enter as a unitary component into a new chunk in a further processing step. This explains why the phenomenal experience changes with practice. But why do the initial primitives evolve into a small number of words instead of innumerable irrelevant processing units?

The reason lies in the combined consideration of two phenomena. The first depends on the properties of the human processing system. The future of the chunk that forms a conscious episode depends on ubiquitous laws of associative learning and memory. If the same experience does not re-occur within some temporal lag, the possibility of a chunk acting as a processing primitive vanishes rapidly, as a consequence both of natural decay and of interference with the processing of similar material. The chunks evolve into primitives only if they are repeated. Thus some primitives emerge through a natural selection process, because forgetting and interference lead the human processing system to select the repeated parts from all of those generated by the initial, presumably mostly irrelevant, chunking of the material. The relevance of this phenomenon becomes clear when viewed in relation to a property inherent to any language. If the speech signal is segmented into small parts on a random basis, these parts have more chance of being repeated if they match a word, or a part of a word, than if they straddle word boundaries (for instance, in the prior sentence, "random" or "basis" have more chance of being repeated elsewhere in the text than "domba"). In consequence, the primitives that emerge from the natural selection due to forgetting and interference are more likely to match a word, or a part of a word, than a between-word segment.

This account has been implemented in a computer program, PARSER (see the Appendix on page 67) and applied to the artificial languages used by Saffran and colleagues. Simulations revealed that PARSER extracted the words of the language well before exhausting the material presented to participants in the Saffran et al. (1996b) experiments. In addition, PARSER was able to simulate the results obtained under attention-disturbing conditions (Saffran et al., 1997) and those collected from 8-month-old infants (Saffran et al., 1996a). Finally, the good performance of PARSER was not limited to the trisyllabic words used by Saffran et al., but also extended to a language consisting of one- to five-syllable words (Perruchet & Vinter, 1998a).



To summarise, we suggest that the discovery of the words results from the interaction between one property of language—essentially that the probability of repeatedly selecting the same group of syllables by chance is higher if these syllables form intra-word rather than between-words components—and the properties of the processing systems—essentially that repeated perceptual chunks evolve into processing primitives, which in turn determine the way further material is perceived.

## GENERALISATION

We have spent a long time considering word discovery in artificial languages because we believe that the basic principles of the explanation proposed in this context extend to the formation of any perceptual and representational units regardless of the domain, the complexity of the material, and the timescale of the learning process. As we have mentioned, PARSER works thanks to the interaction between one property of the language and a few properties of the human processing system. There is no reason to believe that this interaction occurs only with artificial, simplistic languages.

On the one hand, the target property of artificial languages, namely that the probability of repeatedly selecting the same group of syllables by chance is higher if these syllables form intra-word rather than between-words components, is obviously shared by any natural language. Moreover, providing a change from phonological primitives to visual features, the same property is also true for the objects of the real world. For instance, the probability of repeatedly selecting two parts in the environment is stronger if these parts belong to the same object than if they belong to different objects. On the other hand, the properties of the processing system on which PARSER relies are very general.

Note that our solution to the word extraction issue does not involve any new and specialised learning devices. The unitisation of a few primitives, due to their processing within the same attentional focus, is one of the basic tenets of associative learning (Mackintosh, 1975). Besides being necessary, attention is also sufficient for memory and learning to occur. This means that no superimposed operations—such as some forms of intentional orientation towards learning—are required. The resulting picture is that most authors, under different terminology, would acknowledge that learning is an automatic associative process that would associate all the components that are present in the attentional focus at a given point (Frensch & Miner, 1994; Jiménez & Mendez, 1999; Logan & Etherton, 1994; Stadler, 1995; Wagner, 1981). Learning and memory are nothing other than the by-product of attentional processing.

Likewise, the laws of forgetting and the effects of repetition are ubiquitous phenomena. For instance, one fundamental assumption of the model is that a cognitive unit is forgotten when not repeated and strengthened with repetition. This assumption can be taken for granted irrespective of whether the process occurs in the few minutes of an experimental session or across larger timescales, in keeping with a long-standing tradition of research into the laws of memory and associative learning. Moreover, the interdependence of processing units and incoming information—the nature of the processing primitives determines how the material is perceived and the nature of the material determines the transformation of the processing primitives, and so on recursively—is consistent with a developmental principle initially described by Piaget's concepts of assimilation and accommodation (Piaget, 1985). Most current theories of development, although they use different terminology, also rely on the constructive interplay between assimilation-like and accommodation-like processes.

In consequence, PARSER's principles seem to be relevant to finding correct units in natural language and in the other naturalistic domains. Briefly stated, the generality of PARSER is ensured by both the generality of the world property (the most-repeated units are the relevant units of the world) and the generality of the behavioural laws (e.g. only repeated units shape long-lasting representations) on which it relies.

To summarise, the fact that conscious percepts are capable of self-organisation, initially applied to the word extraction issue in laboratory situations, suggests a new account of the human ability to build internal representations isomorphic to the actual world units.[3] In this account, conscious representations participate in their own development. Basic principles of associative learning and memory allow conscious representations to reach their high degree of organisation and adaptiveness, provided we consider that associations occur between the rich content of conscious experiences. The notion of self-organisation excludes any organising cognitive systems or principles that would be superimposed on the phenomenal consciousness. The phenomenal consciousness itself ensures its own improvement in representational power, thanks to the propensity of conscious representations to evolve in accordance with basic associative learning principles.

[3] Of course, this isomorphism is not perfect. Firstly, the representations we create are limited by sensory constraints. For instance, we do not have any perception about the sounds outside the 20–20,000 Hz range, or about the light sources in the infrared or ultraviolet wavelength range. Likewise, phenomenal experience does not provide us with any direct representation of the structure of the physical world at other scales, such as the atomic microstructure or galactic organisation. Secondly, even the parts of the world directly available to our sensory equipment can be misrepresented. For instance, in the perceptual illusions, perceptual processes generally well suited in natural situations cease doing their job reliably when faced with very special patterns. However, for the sake of brevity, we go on speaking hereafter about the isomorphism between subject's representations and world structure, even though the very phenomenon the expression recovers cannot be described as a simple term-to-term matching.



## RETHINKING IMPLICIT LEARNING

Even if one admits that the above account works well for the formation of conscious representations of parts of the world, such as words and objects, it turns out that those aspects are not usually conceived of as directly related to implicit learning. More generally, the existence of linguistically or physically relevant representations is not commonly considered as sufficient for accounting for human mental activities. Representations are generally construed as the elementary building blocks of thought, and mind activities are assumed to include the formation of knowledge in which the blocks are combined on the basis of some organising (e.g. logical, syntactical, or statistical) principles.

Our proposal is that the notion of self-organising consciousness offers a way of thinking about these complex aspects of behaviour without having recourse to the notion of unconscious rule abstraction or unconscious statistical computation. The idea is that the separation between basic units on the one hand, and rules governing those units on the other, or between lexicon and syntax in linguistic terminology, is warranted in a scientific approach (i.e. from the observer's viewpoint) but, at least on some occasions, might be irrelevant for the processing system. The purpose of the processing system is to generate a representation of the world that integrates the whole ongoing information (internal and external) into a coherent and meaningful event. This complex and integrative representation, we argue, makes any other forms of knowledge or computation devoid of any object. This thesis relies heavily on the idea that neural systems "trade representation against computation", to borrow the expression coined by Clark and Thornton (1997).

In Perruchet and Vinter (unpublished manuscript), we discuss at length how conscious representations can account for improved performances in various rule-governed situations. We argue that all the cases in which analytical operations, such as hypothesis testing, are apparently performed by the mind in the absence of any form of conscious awareness can be encompassed within the notion of self-organising consciousness. We saw earlier how the notion of self-organising consciousness allows us to account for the formation of internal representations that are increasingly congruent with the world structure. If we expand the scope of these representations to the various dimensions involved in a given problem, it becomes conceivable that a representation contains, in some sense, both the data and the solution of the problem. The solution pops up in the mind because it is a part of the model of the world that people have built through automatic associative processes. Let us take a simple example, one relating to the notion of transitivity. In the linear ordering tasks, two premises are presented, the formal expression of them being: A is longer than B and B is longer than C. Participants have to judge whether an expression such as: A is longer than C is correct. It can be assumed that people solve this task because they have some formal notion about the transitivity of the expression "larger than", and that they apply the transitivity rule to the problem at hand. However, people could also have built an integrative representation of the premises in the form of a linear array, and then read the response to the question directly from this representation. There is now a consensus about the idea that people proceed in this way (Evans, Newstead, & Byrne, 1993). This illustrates how a representation which is isomorphic to the world structure makes rule knowledge unnecessary.

How does our account apply to traditional situations of implicit learning? Our re-interpretation (e.g. Perruchet & Vinter, 1998b; Perruchet, Vinter, & Gallego, 1997) is that the training phase modifies the way the data are consciously coded and perceived. Let us consider artificial grammar learning as an illustration. Assuming that, say, XRX is a frequent recursion in the set of strings generated by a finite-state grammar, participants no longer perceive X and R as two familiar but separate entities, but perceive XRX as an increasingly familiar unit. One possible explanation for the above-chance grammaticality judgements of a new string including XRX is that participants interpret, more or less automatically, the level of perceptual fluency as an indicator of grammaticality. Strings that can be read easily because chunks of letters are directly perceived as familiar units would tend to be judged as grammatical. In short, in our re-appraisal, the formation of the conscious unit XRX replaces the unconscious extraction, retention, and use of a rule such as: If XR, then X.

It might seem, at first glance, that any fragment of a grammatical utterance is itself grammatical, and can be recombined with another fragment to form a new grammatical string. Given this logic, the initial chunking of the material would not matter. And indeed, the notion of "fragmentary knowledge" conveys the tacit implication that it is a quite impoverished form of knowledge. This view is faulty, as can be illustrated using the example of natural language. For instance, in the preceding sentence, the segments: "this view", or "natural language" form structurally relevant sequences, in the sense that they can be recombined with a large number of other sequences, whereas "faulty, as can" cannot be easily integrated as a component in another linguistic context, although it is a component of a legal sentence. It is obvious that it is preferable to become familiar with the former segments than with the latter. Likewise, in the letter strings generated by a finite-state grammar, it is preferable to become familiar with a subset of fragments—for instance, those that are generated by a recursive loop—than with other, randomly selected, fragments. We (Perruchet, Vinter, Pacteau, & Gallego, in press) have shown that participants in an artificial grammar learning setting indeed formed the structurally relevant units. They were asked to read each string generated by a finite-state grammar and, immediately after reading, to mark with a slash bar (/) the natural segmentation positions. The participants repeated this task after a phase of familiarisation with the material, which consisted of learning items by rote, performing a short-term matching task, or searching for rules. The same number of total units was observed before and after the training phase, thus indicating that participants did not tend to form increasingly larger units. However, the number of different units reliably decreased, whatever the task during training. This result was taken as evidence that participants' processing units became increasingly relevant as training progressed (see also Servan-Schreiber & Anderson, 1990). Perruchet et al. (in press) also showed that PARSER, the computer model that was used previously to account for the discovery of words in an unsegmented speech flow (Perruchet & Vinter, 1998a; see p. 67), also accounted for participants' actual performance. Thus the principles that make it possible to discover the lexical units of an artificial language built from the random concatenation of words also proved to be efficient in the discovery of the syntactically relevant units of an artificial language built from a finite-state grammar.



It is worth examining why such simple principles work well in a situation that was once thought of as involving grammatical rule abstraction. It is because first-order and second-order dependency rules capture virtually all the structural constraints of the standard finite-state grammars. For instance, Perruchet and Gallego (1997) have demonstrated that consideration of only the first-order dependency rules is sufficient to account for the performance of the participants in the Reber (1976) experiments and many others that use the same material.

The same demonstration can be repeated for other standard situations of implicit learning. Regarding the repeated sequence tasks, the situation is still simpler. Indeed, because the same sequence is repeated continuously, the way this sequence is segmented is of no consequence: no chunk is any more relevant than any other. There is a prerequisite, nevertheless: Namely, that the same parsing occurs throughout the session. Empirical data support the point. Of special interest in this context is the Stadler (1995) study. Stadler noted that the tone-counting task used in sequence learning paradigms can alter performance by dividing a pool of limited resources—the conventional interpretation of attentional effects—but also by disrupting the organisation of the sequence. Indeed, although a tone occurs after each trial, only some tones (e.g. high-speech tones) trigger the update of the running count. These tones, presumably, serve as boundaries for successive perceptual chunks. If these tones are introduced randomly then the content of the perceptual chunks changes continuously throughout the session, a circumstance that might impair learning. Stadler has shown that this hypothesis is correct: a task preventing the repeated processing of the same chunks is detrimental for learning, even though overall resources are left intact, whereas a task with additional attentional demands respecting the perceptual organisation has no effect. Stadler's study provides a strong indication that the function of attention must not be understood only as a pool of unspecific resources, and that the content of successive attentional focuses, or in other words the subjects' phenomenal experience of the task, is a determinant of performance improvement.

## ABSTRACTING AWAY FROM THE SENSORY CONTENT

In the preceding section, we claimed that the changes in the way we consciously perceive and represent our environment might underlie some apparent phenomena of syntax sensitivity. In some cases, it is easy to see how a simple representation could replace genuine rule knowledge. For instance, it is easy to see how perceiving XRX as a unit could replace the rule: "If XR then X". However, adaptation to other situations does not seem reducible to the same approach. An especially striking example is provided by the studies that reveal participants' ability to abstract away from the sensory content of the training situation, an ability that cannot seemingly be explained by any association-based account.

### Experimental evidence for abstraction

As a case in point, let us consider the recent experiments by Marcus, Vijayan, Rao, and Vishton (1999). Seven-month-old infants were exposed to a simplified, artificial language during a training phase. Then they were presented with a few test items, some of which belonged to the same language whereas others introduced some structural novelty. The infants controlled the exposure duration of the stimuli by their visual fixation on a light. Their discrimination was assessed through their longer fixation (and hence listening) times for items introducing structural novelty.



The point of interest is that Marcus and co-workers introduced a change in the sensory content of the material between the training and the test phases. For instance, in one experiment, infants heard 16 three-word sentences such as *gatiti*, *linana*, or *tanana*, during the training phase. All of these sentences were constructed on the basis of an ABB grammar. The infants were then presented with 12 other three-word sentences, such as *wofefe* and *wofewo*. The crucial point is that although all of the test items were composed of new words, only half of them were constructed from the grammar with which the infants had been familiarised. In the selected example, the grammatical item was *wofefe*. *Wofewo* introduces a structural novelty in that it is generated from a concurrent ABA grammar. The infants tended to listen more to the sentences generated by the ABA grammar, thus indicating their sensitivity to the structural novelty. In another experiment, infants were shown to be able to discriminate sentences generated by an AAB grammar.

Similar studies using more complex material have been performed with 11-month-old infants (Gomez & Gerken, 1999) and with adults. In some studies involving artificial grammar, the letters forming the study items are changed in a consistent way for the test of grammaticality (e.g. C is always replaced by X, B by L, and so on). Reber (1969), and several subsequent studies (Dienes & Altman, 1997; Manza & Reber, 1997; Mathews et al. 1989; Shanks, Johnstone, & Staggs, 1997; Whittlesea & Wright, 1997) have shown that participants still outperform chance level under these conditions. The principle underlying the transfer in the so-called "changed letter procedure" has been extended to other surface changes. For instance, the training items and the test items might be, respectively, auditory items and visual items (Manza & Reber, 1997), colour and colour names, sounds and letters (Dienes & Altman, 1997), or vice versa. Successful transfer was observed in each case. Reber claimed that these results testify to the fact that participants are able to abstract the "syntax" of the displayed material, independently of the "vocabulary".

At first glance, evidence for transfer between event patterns cutting across their sensory contents cannot be accounted for by our model of implicit learning. Indeed, the formation of an associative link between, say, *ga*, *ti*, and *ti*, whatever its strength, seems fundamentally unable to explain transfer to *wo*, *fe*, and *fe*, as observed in the Marcus et al. (1999) experiments. Accordingly, Marcus et al. concluded that infants have the capacity to represent algebra-like rules and, in addition, "have the ability to extract those rules rapidly from small amounts of input and to generalise those rules to novel instances" (p. 79). Likewise, Reber (1993), talking about performance in the transfer letter paradigm in artificial grammar learning studies, claimed that "the abstractive perspective is the only model of mental representation that can deal with the existence of transfer of knowledge across stimulus domains" (Reber, 1993, p. 121).

## The outline of a re-appraisal

We have no problem with the claim that the evidence of transfer reviewed above is indicative of abstraction. However, we challenge the view that abstraction is indicative of rule formation and rule use. Our claim is that transfer is a natural implication of the SOC model.

Let us return to PARSER. PARSER shows how the initial conscious percept, which is generally irrelevant to the material structure, becomes increasingly isomorphic with the structurally relevant units, thanks to the elementary principles of associative learning and memory. On pages 46–49, we considered that the initial percept exactly matched the content of the perceived stimuli. For instance, given the auditory string *badubatibu*, we assume that participants first form the auditory units *baduba*, *tibu*, and so on, by chunking together the auditory primitives *ba*, *du*, *ti*, and *bu*, and this assumption was sufficient to account for the data. However, it is worth stressing that this assumption is notoriously restrictive. Indeed, the primitives that enter into the associations are internal representations that only partially match the external stimuli that trigger these representations. For instance, as a result of earlier associations, the representations of *ba*, *du*, *ti*, and *bu*, involve a written component in literate people. Thus, when a new association is built between, say, the components of the auditory percept *baduba*, the new unit is not limited to the auditory domain, but naturally extends to the area of generalisation of the primitive components, and especially to the visual domain. More generally, many examples of transfer originate in the fact that conscious primitives entering into the new associations are not tied to a fixed, domain-specific format of representation, but are instead often amodal, flexible, and domain-general. Conscious knowledge is represented into a cross-system code (e.g. Fodor, 1983), a property that ensures that any conscious content possesses a certain abstractness.

Going a step further, it can also be argued that when a few syllables are perceived within one attentional focus, the resulting conscious experience is not necessarily limited to the sum of these syllables (even considering that they are represented into a cross-system code) but instead could embed some direct perception of the overall structure. For instance, *baduti* will not be perceived as *bababa* or *baduba.* The obvious difference lies in the number of repetition of the same primitives. There is no doubt that a part of the representation of *bababa* is that it consists in the repetition of the same syllable (a "run"), and that a part of the representation of *baduba* is that the same syllable is repeated with an intervening syllable (a "trill"). Coding a pattern as a run or a trill entails some form of relational coding, the relation involved here being the same–different relationship. Thus, our assumption is that the sensory input processed within one attentional focus can also integrate some relational information.

If we take it for granted that such abstract and relational features are parts of conscious representations, then there is no reason not to apply to these features the same reasoning that we applied to PARSER's primitives. Abstract features, if they are frequently involved in the conscious perception of a given material, can emerge from noise on the basis of a selection process analogous to the one that we showed to be responsible for the formation of specific representations. As is the case for specific representations, the extraction of regularities is facilitated by the fact that, in its turn, the initial perception determines the way further material is perceived. Thus, when some abstract relations have been perceived frequently enough to become perceptual primitives, they are automatically detected in the new material whenever present. However, in this case, the end-product of the process will be the emergence of representations coding the deep structure of the situation at hand, which makes transfer to other surface features natural. To oversimplify the matter for the sake of understanding, one could say that, in the conventional account, perception provides the system with a database composed of elementary, specific primitives, from which the unconscious processor abstracts the deep underpinning rules. In our account, the primitives are a little more abstract and complex features. However, with these new primitive units, no further conceptual operations are needed to account for transfer.

It is worthy of note that this interpretation is viable only if the coding of the incoming information in an abstract and relational format remains simple enough to be attributed to low-level perceptual processes. Admittedly, if it turns out that the perceptual primitives needed to account for the available data are, say, nested high-level order dependency rules, it would be unrealistic to claim that these primitives are directly coded by elementary perceptual mechanisms. Thus it is important to show that the available evidence of transfer can be explained in terms of the coding of fairly simple relations. In the following section, we examine the form of abstract and relational coding needed to account for the available findings on transfer. We will show that only surprisingly simple forms of coding are required.

## Perceptual primitives can be abstract and relational

To begin with the most simple case, let us consider the Manza and Reber (1997) results, showing a transfer between auditory and visual modalities in the artificial grammar learning area. These authors interpret their findings as providing support for their abstractionist, rule-based view. Although the authors do not make their interpretation more explicit, we assume that their line of reasoning could be as follows. If, for instance, subjects perceive the visual sequence XMX, they abstract the knowledge that the letter X can be repeated with a lag of one letter. When they again perceive XMX, but in the auditory modality, they might experience some familiarity with the display, because the same rule applies. This interpretation undoubtedly works well. However, the phenomenon can be explained easily without having recourse to rules. It suffices to consider that there is a direct correspondence between the visual and the auditory format of the letters X and M. It is worth stressing the differences between the two approaches. In the former case, a rule-governed pattern needs to be extracted from the visual stimuli, before being transferred to the auditory stimuli. In the latter case, matching is direct and independent of the structure of the material. A simple thought experiment can help to clarify the differences and, by the same token, demonstrates the irrelevance of a rule-based account. Suppose that the material is generated randomly, instead of by a finite-state grammar, and thus presents no rule-governed, salient pattern. For the sake of illustration, suppose that a string such as XMT is presented. In a rule-based interpretation, transfer should not occur, because a structure cannot be abstracted. Now, it is quite obvious that the prior auditory presentation of XMT increases familiarity with the visual display XMT even though there is no common salient structure.

The same comment can be applied to some other studies. For example, Dienes and Altman (1997) observed a positive transfer between colours and the name of colours, which can also be accounted for by the natural mapping between the primitives involved in the experiment. Again, transfer would probably occur even with randomly generated stimulus sequences, thus demonstrating the irrelevance of a rule-based interpretation. However, not all studies of transfer can be explained using so simple an argument. As a case in point, the above explanation does not apply to the Marcus et al. study in which transfer is observed between, say, *gatiti* and *wofefe*, because there is no natural mapping between *ga* and *wo*, or *ti* and *fe*.

Re-interpretation of the Marcus et al. data demands recourse to another property of conscious percepts, namely the direct coding of simple relations between the components of one percept. The relation that needs to be coded is the relation "same–different", or, in other words, the only ability that infants need to exhibit is that of coding the repetition of an event. If one postulates that infants are able to detect whether two successive stimuli are the same or not, the Marcus et al. results are easily explained. Indeed, as pointed out by McClelland and Plaut (1999), *gatiti*, *wofefe*, and, more generally, all the ABB items, can be coded as different–same, whereas none



of the other items can be coded using the same schema. AAB items are coded as same–different; ABA items instantiate a slightly more sophisticated pattern. Note that there is no indication in the data that this pattern is actually perceived as special: Considering that ABA items do not match the pattern of the other items is sufficient to account for the data. However, it does not seem to be unrealistic to assume that a trill pattern is also directly perceived when the components of this pattern can be processed within a single attentional focus. The numerous studies showing infants' early sensitivity to symmetrical displays support this assumption.

At first glance, the demonstrations of transfer stemming from the more complex situations of artificial grammar learning in adults imply the coding of far more complex relations. We now argue that in fact, as surprising as this conclusion might be, the very same abilities that we have invoked up to now are sufficient. Indeed, although finite-state grammars embed complex relations, the coding of fairly simple patterns appears sufficient to account for improved performance in transfer situations. For instance, Whittlesea and Wright (1997, Experiment 4) reported successful transfer between letters and colours in artificial grammar learning. In their experiment 4, 5 out of the 20 training items begin with a salient alternation ("RMR"). Now, it turned out that colour alternation at the beginning of a string appeared in legal test items but never in illegal test items. It is enough to assume that participants consider the test items beginning with an alternation to be grammatical, and respond at random on the others, to simulate observed performance. If we take this interpretation for granted then transfer is easy to account for. Indeed, although there is no natural link between, for instance, R and a red square, a natural mapping may be established between the subjective unit "RMR" and "RED/YELLOW/RED", or any other colour alternation. Again, the observation of a positive transfer is irrelevant as to whether subjects have abstracted the complex grammar used to generate the material. It can be accounted for more parsimoniously by assuming that subjective units are at least partially represented into a relational code.

For a still more complex illustration, let us consider one of the recent studies by Shanks et al. (1997), which concluded that transfer in artificial grammar learning is mediated at least to some extent by abstract knowledge. Experiment 1 used a standard changed-letter procedure, in which the letters used during study—M, R, T, V, and X—were replaced by C, H, J, L, and N, respectively, for the test. Shanks et al. introduced five types of violation in their ungrammatical transfer strings. The only violation that led participants to reject the strings in a forced choice grammaticality test was illegal letter repetitions. In the original grammar, only R, T, and V could be repeated. Thus, in legal transfer items, H, J, and L could also be repeated, but C and N could not. Shanks et al. showed that participants rejected transfer items including a repetition of one of these two letters at a significant level. Such a result suggests that subjects were able to perform a quite sophisticated analysis, including at least two steps. They first have to identify the fact that M and X were never repeated in the original set, then to establish a correct mapping between M and C, on the one hand, and X and N on the other.

It can be shown that correct responses imply neither of these steps. Let us assume that participants have formed subjective units, each composed of a few letters. An examination of the training strings shows that these subjective units include far fewer repetitions than if letters had been selected at random. The training strings included nine repetitions, whereas we assessed (through a computational simulation) the number of repetitions expected by chance at about 22. Now, looking at the five pairs of transfer strings testing the "illegal letter repetition" feature, it appears that ungrammatical test strings always include more letter repetitions than grammatical test strings. It is enough for the participants to feel the encoding units including a letter repetition to be unfamiliar for them to choose the grammatical item from each pair. The point is that there is strictly no need to infer what letter repetitions were legal in the study strings, or to establish a letter-to-letter mapping: it suffices to be sensitive to the fact that subjective units rarely include a letter repetition, whatever the nature of these letters. Transfer originates in the fact that a unit's feature, such as "including a letter repetition", can be captured naturally, and not in the abstraction of the rules of the finite-state grammar used to generate the letter strings (for other analyses pointing out to the primary importance of repetition structure to account for transfer in artificial grammar learning, see Gomez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 1999).

## Is our account of transfer more parsimonious?

To recapitulate, in the conventional models, the data made available to the central processor are the individual sensory-based events. The task of finding analogies between events that differ in their surface appearance is the job of some further inferential processes. These processes belong to the domain of cognition and, more precisely (because we are not aware of them), to the realm of the sophisticated cognitive unconscious. In our alternative conception, unconscious (but elementary) processes provide a conscious representation of the sensory input that is framed directly in some abstract and relational way, as any conscious content is. With this modified input, the performance observed in transfer situations no longer needs to be explained in terms of a sophisticated unconscious processor. The ubiquitous learning and memory processes evoked in the previous sections are sufficient to explain the emergence of a reliable representation of the deep structure of the material. We have indicated how simple principles of associative learning and memory explain the emergence of conscious representations that are increasingly isomorphic to the world structure in cases where the sensory domain remains identical. When applied to more abstract primitives, the very same principles account for the discovery of the structure of the material in cases where the sensory domain is changed.



Opponents of this position might argue that our conception simply shadows or resituates the problem instead of solving it. The argument should be that positing that ongoing sensory information is directly coded into an abstract and relational code is akin to taking as a premise the to-be-explained phenomenon (i.e. the ability to transfer), and presumably further consideration of this initial stage of processing would indicate that it, in fact, involves the same kind of complex machinery that most authors include under the label of Cognitive Unconscious. This criticism is unsound, however, because the relationships we assume to be coded directly by low-level perceptual processes are considerably simpler than the abstract rules of the mainstream tradition. They are limited to a few aspects, including the same–different distinction, the properties of symmetry, repetition, and alternation, and relationships along some perceptual dimensions such as "smaller than" or "brighter than". It is not biologically implausible to assume that these relationships are coded at earlier stages of neural processing, although there is as yet no direct evidence (one exception is the direct coding of the relation "brighter than", that is at least partially coded at the retinal level by lateral interaction between concurrent stimulations).

In the absence of more extensive neuropsychological arguments, our hypothesis finds some support in the primacy of relational coding in phylogenetic evolution. It has long been shown that animals such as rats are able to perform tasks involving elementary forms of relational learning. For instance, if rats are trained with two stimuli differing in brightness in such a way that the choice of the brighter is rewarded and the choice of the darker is not rewarded, they subsequently choose the brighter of two new stimuli, even though the absolute brightness of the new rewarded stimulus might be identical to that of the old unrewarded stimuli. Thus rats appear to be sensitive to the relationship between stimuli rather than to their absolute properties. Such a demonstration has been replicated with various animal species and using a variety of simple relationships, such as "larger than". Primates and a number of birds also appear able to learn a discriminative response to pairs of stimuli, depending on whether they are identical or different and, once acquired, this ability can transfer to any new stimulus pair, irrespective of its nature. Within the perspective of evolutionary biology, these results are not at all surprising. In many cases, the raw information provided by an isolated event is only partially relevant. For instance, the retinal size of a perceived object or animal is uninformative, because it depends on the distance between the observer and the distal stimulus. Similarly, the absolute brightness provides incomplete information because perceived brightness depends on the ambient luminance. Considerably more reliable information is provided by a relational coding by means of which the size or brightness of a new stimulus is assessed by comparison with contextual stimuli.

## CONSCIOUSNESS: FROM "NECESSITY" TO "SUFFICIENCY"

This chapter has proposed a new view of implicit learning, in which the phenomenon is conceived of as a transformation of conscious experiences through the action of elementary associative processes acting on the components of these experiences. To conclude, we would like to come back on the deep theoretical commitments of this view, and briefly to address some of the most common objections to it.

Our framework provides a specification of a meta-theory of the mind, partly rooted in the philosophical approach developed by Searle (1992), and, on the psychological side, in the work of Dulany (1997), who called it, for want of a better term, the "mentalistic" framework. The most salient feature of the mentalistic framework is the denial of the Cognitive Unconscious. The core of this position is the refutation of the notion of unconscious representations. Of course, this position also makes it necessary to reject the possibility of performing unconscious manipulations and transformations of these representations and, hence, the notions of unconscious rule abstraction, computation, analysis, reasoning, and inference, become meaningless. The only representations people create and manipulate are those that form the momentary phenomenal experience. To quote ourselves: "Processes and mechanisms responsible for the elaboration of knowledge are intrinsically unconscious, and the resulting mental representations and knowledge are intrinsically conscious. No other components are needed" (Perruchet et al., 1997, p. 44).

Several criticisms can be levelled at such a view. Some of them stress that there are several well-known empirical phenomena, such as the subliminal semantic activation effect, that provide direct demonstrations of the cognitive unconscious. The reliability of these alleged demonstrations has been questioned in earlier papers (Holender, 1986, Shanks & St John, 1994: see also Perruchet & Vinter, unpublished manuscript), and we do not deal further with this issue here. Other potential criticisms are directed more against the *a priori* interest or legitimacy of the approach than against its empirical likelihood. Arguments start from the idea that the main object of psychological science is the processing of information, such as it can be described in a program for a digital computer. Now, consciousness by itself is a property that is computationally irrelevant for information processing



modelling. The functioning of any model does not depend in any way on the conscious/unconscious status we ascribe to the representations it manipulates and to the operations it performs. As a consequence, the question of knowing whether consciousness is necessary for any cognitive activity is not scientifically warranted, or, at best, needs to be postponed until substantial progress has been made regarding the "serious issues".

We acknowledge that consciousness, in itself, is computationally irrelevant. By the way, we also acknowledge that, because cognitive activities can be simulated on a machine, one is unable to prove the necessity of consciousness when cognitive activities are implemented in a brain. This does not mean that the issue of consciousness is irrelevant in a scientific approach. Indeed, positing that the model has to simulate conscious states while respecting the properties of conscious thought introduces considerable structural and functional constraints for the model. Attention and consciousness are linked to the notions of limited capacity, seriality and relative slowness of processing, as well as quick memory decay. These constraints are relevant for an information processing approach. They can be implemented in a computational model to address the following question: Is a model that is fulfilling these constraints sufficient to account for human thought and behaviour?

To illustrate, let us consider PARSER, the computational model of word segmentation that we present as supportive of our approach. Of course, PARSER is not conscious, nor are the computers on which it runs. The model would work as well whether the representations it manipulates are assigned a conscious or unconscious status. The relevance of PARSER with regard to our concern lies elsewhere. It lies in the fact that PARSER respects the striking constraints inherent to conscious thought. Thus the only representations included in the model closely match the conscious representations subjects might have when performing the task. The early coding of the material as a set of short and disjunctive units, as well as the final coding of the input as a sequence of words, are assumed to closely match the phenomenal perceptual experience of listeners. This correspondence also extends to the entire learning phase, thus permitting our model to perform word segmentation while mimicking the on-line conscious processing of incoming information. By doing so, PARSER demonstrates that conscious percepts and representations are sufficient for word extraction.

Generalising from word extraction along the lines set out in this chapter, our claim is that conscious percepts and representations are sufficient, from a computational standpoint, to account for the adaptive abilities described in the implicit learning literature, including those that seemingly point to rule abstraction. Here is, we believe, the ultimate interest of our approach: accounting for complex non-intentional learning phenomena without relying on the unlimited abilities of a putative cognitive unconscious.

## ACKNOWLEDGEMENTS

This work has been supported by the Centre National de la Recherche Scientifique (CNRS, UMR SO22), the Universite de Bourgogne, and the Region de Bourgogne (AAFE).

## REFERENCES

Berry, D.C., & Broadbent, D.E. (1988). Interactive tasks and the implicit–explicit distinction. *British Journal of Psychology*, *79*, 251–272.

Clark, A., & Thornton, C. (1997). Trading spaces: Computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*, *20*, 57–90.

Cleeremans, A. (1993). *Mechanims of implicit learning: A connectionnist model of sequence processing*. Cambridge, MA: MIT Press/Bradford Books.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406–416.

Cohen, A., Ivry, R.I., & Keele, S.W. (1990). Attention and structure in sequence learning, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 17–30.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Crabb, B.T., & Dark, V. (1999). Perceptual implicit memory requires attentional encoding. *Memory and Cognition*, *27*, 267–275.

Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* Oxford: Oxford University Press.

Dienes, Z., Broadbent, D., & Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 875–887.

Dulany, D.E. (1997). Consciousness in the explicit (deliberative) and implicit (evocative). In J.D. Cohen & J.W. Schooler (Eds.), *Scientific approaches to the study of consciousness* (pp. 179–212). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Evans, J.St.B., Newstead, S.E., & Byrne, P. (1993). *Human reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Fisk, A.D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 181–197.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Frensch, P.A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 567–584.

Frensch, P.A., & Miner, C.S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory and Cognition*, *22*, 95–110.

Ganor-Stern, D., Seamon, J.G., & Carrasco, M. (1998). The role of attention and study time in explicit and implicit memory for unfamiliar visual stimuli. *Memory and Cognition*, *26*, 1187–1195.

Gomez, R.L., & Gerken, L.A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135.

Gomez, R.L., Gerken, L.A., & Schvaneveldt, R.W. (2000). The basis of transfer in artificial grammar learning, *Memory and Cognition*, *28*, 253–263.



Green, R.E.A., & Shanks, D. R. (1993). On the existence of independent learning systems: An examination of some evidence. *Memory and Cognition*, *21*, 304–317.

Hayes, N.A., & Broadbent, D. (1988). Two models of learning for interactive tasks. *Cognition*, *28*, 249–276.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1–23.

Hsiao, A.T., & Reber, A. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495–531). Thousand Oaks, CA: Sage Publications.

Jiménez, L., & Mendez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 236–259.

Logan, G.D., & Etherton, J.L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1022–1050.

McClelland, J.L., & Plaut, D.C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Science*, *3*, 166–168.

MacDonald, P., & MacLeod, C.M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, *98*, 298–310.

McGeorge, P., & Burton, M. (1990). Semantic processing in an incidental learning task. *Quarterly Journal of Experimental Psychology*, *42*, 597–609.

Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.

Mandler, G. (1975). Consciousness: Respectable, useful, and probably necessary. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 229–254). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Manza, L., & Reber. A.S. (1997). Representing artificial grammar: Transfer across stimulus forms and modalities. In D. Berry (Ed.), *How implicit is implicit learning?* Oxford: Oxford University Press.

Marcus, G.F., Vijayan, S., Rao, S.B., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

Mathews, R.C., Buss, R.R., Stanley, W.B., Blanchard-Fields, F., Cho, J.-R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1083–1100.

Miller, G.A. (1962). *Psychology: The science of mental life*. New York, Harper & Row.

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

O'Brien, G., & Opie, J. (1999). Putting content into a vehicle theory of consciousness. *Behavioral and Brain Sciences*, *22*, 175–196.

Perruchet, P., & Amorim, M.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.

Perruchet, P., & Gallego, J. (1997). A subjective unit formation account of implicit learning. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 124–161). Oxford: Oxford University Press.

Perruchet, P., & Vinter, A. (1998a). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*, 246–263.

Perruchet, P., & Vinter, A. (1998b) Learning and development: The implicit knowledge assumption reconsidered. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495–531). Thousand Oaks, CA: Sage Publications.

Perruchet, P., Vinter, A., & Gallego, J. (1997). Implicit learning shapes new conscious percepts and representations. *Psychonomic Bulletin and Review*, *4*, 43–48.
Perruchet, P., Vinter, A., Pacteau, C., & Gallego, J. (in press). The formation of structurally relevant units in artificial grammar learning. *Quarterly Journal of Experimental Psychology.*
Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development.* Chicago: University of Chicago Press.
Posner, M.I., & Boies, S.J. (1971). Components of attention. *Psychological Review*, *78*, 391–408.
Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, *81*, 115–119.
Reber, A.S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.
Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious.* New York: Oxford University Press.
Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 275–285.
Roediger, H.L., III (1990). Implicit memory: A commentary. *Bulletin of the Psychonomic Society*, *28*, 373–380.
Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996b). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996a). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.
Saffran, J.R., Newport, E.L., Aslin, R.N., Tunick, R.A., & Barrueco, S. (1997). Incidental language learning. *Psychological Science*, *8*, 101–105.
Schmidt, P.A., & Dark, V.J. (1998). Attentional processing of "unattended" flankers: Evidence for a failure of selective attention. *Perception and Psychophysics*, *60*, 227–238.
Searle, J.R. (1992). *The rediscovery of the mind.* Cambridge, MA: MIT Press.
Servan-Schreiber, D., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592–608.
Shanks, D.R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1435–1451.
Shanks, D.R., Johnstone, T., & Staggs, L. (1997). Abstraction processes in artificial grammar learning. *Quarterly Journal of Experimental Psychology*, *50A*, 216–252.
Shanks, D.R., & St John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367–447.
Stadler, M.A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 674–685.
Tunney, R.J., & Altmann, G.T.M. (1999). The transfer effect in artificial grammar learning: Re-appraising the evidence of transfer of sequential dependencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1322–1333.
Velmans, M. (1999). Neural activation, information, and phenomenal consciousness. *Behavioral and Brain Sciences*, *22*, 172–173.
Wagner, A.R. (1981). SOP: A model of automatic memory processing in animal behaviour. In N.E. Spear & R.R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
Whittlesea, B.W.A., & Wright, R.L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 181–200.



## APPENDIX: PARSER

PARSER is centred on a single vector, called Percept Shaper (PS). At the start, PS contains only the primitives composing the material, namely a few syllables. Learning proceeds through the iterative processing of small successive parts of the linguistic corpus. Each part is composed of one to three processing primitives (the number is determined randomly for each percept), thus simulating the successive attentional focuses of a human subject processing the same corpus. Each perceived part is added to PS, and can itself serve as a new primitive for the shaping of subsequent inputs, as the syllables did initially. This simulates the fact that perceptual contents are changing throughout the task. Finally, if learning has been successful, PS contains all the words, and only the words of the language.

Why does PS not become encumbered with an innumerable set of irrelevant and increasingly lengthy units? It is because the future of a unit depends on its weight, which represents trace strength. The weight of a given unit is incremented each time this unit is perceived (weight = +1), and decremented each time another unit is perceived (decrement = –0.05). Decrement simulates forgetting (in the original program there was also some interference, the computational details of which are irrelevant here). To fulfill its shaping function, any unit of PS needs to reach a threshold value (threshold = 1). As a consequence, a unit needs to be perceived repeatedly and regularly to persist on fulfilling a shaping function. In contrast, when the frequency of perception of a given element is not high enough to counteract the effects of forgetting and interference, this element is removed from PS when its weight becomes zero.

This simplistic algorithm is sufficient for PARSER to achieve the extraction of the words forming the artificial languages designed by Saffran and colleagues, with a much more limited amount of practice than real subjects need (Perruchet & Vinter, 1998a). However, it must be understood that the details of the functioning of the model are not intended to provide a realistic picture of the processes that are actually involved. As a case in point, forgetting is simulated through a linear decrement, whereas there is evidence that the forgetting curve fits only moderately well with a linear trend. A more recent version includes a more realistic power function.

More importantly, it can be argued that the general architecture of PARSER is not compatible with the meta-theory of the mind underlying this chapter. Indeed, PS can be thought of as a memory store or a mental lexicon, in which symbolic representations are assumed to be potentially active independently of the current phenomenal experience of the subject. This possibility is actually not allowed in our general framework. The contradiction is indeed patent, but, we believe, not detrimental to the demonstration provided by PARSER of the power of the general principles it implements. Indeed, the representations stored in PS play a role only when they match the external input. They perform no function apart from shaping the momentary percepts, that is to say when they enter as a component of the current phenomenal experience. As a consequence, the same result should have been obtained had the memory of the system been simulated as a capacity to build an on-line representation in the presence of a given input, without directly storing the representation itself. In fact, neural network modelling should certainly have been more in the spirit of our approach because it naturally implements the idea that the memory of the system is not necessarily a list of symbolic tokens. However, current connectionist models have fixed input units, whereas a learning principle constitutive of our approach is that percepts evolve throughout training. PARSER is indeed formal in nature, but this choice entails no allegiance to the principles usually found to underpin formal artificial intelligence.