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

PSYCHOLOGY PRESS
ALERE FLAMMAM
Taylor & Francis Group

Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235.
Reber, A.S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition, 5*, 333–361.
Redington, M. (1998). *On hybrid models of artificial grammar learning: Commentary on Knowlton and Squire (1996) and Meulemans and Van der Linden (1997)*. Unpublished manuscript.
Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General, 125*, 123–138.
Redington, M., & Chater, N. (1998). *Computational models of artificial grammar learning*. Unpublished manuscript.
Servan-Schreiber, E., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 592–608.
Shanks, D.R. (1997). Representation of categories and concepts in memory. In M.A. Conway (Ed.), *Cognitive models of memory* (pp. 111–146). Hove: UK: Psychology Press.
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.
Squire, L.R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44*, 453–495.
St John, M.F., & Shanks, D.R. (1997). Implicit learning from an information processing standpoint. In D.C. Berry (Ed.), *How implicit is implicit learning?* (pp. 162–194). Oxford: Oxford University Press.
Tulving, E. (1983). *Elements of episodic memory*. Oxford: Oxford University Press.
Turner, C.W., & Fischler, I.S. (1993). Speeded tests of implicit knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1165–1177.
Vokey, J.R., & Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 328–344.
Vokey, J.R., & Brooks, L.R. (1994). Fragmentary knowledge and the processing-specific control of structural sensitivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1504–1510.
Whittlesea, B.W.A. (1997a). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. In D.L. Medin (Ed.), *The psychology of learning and motivation* (vol. 37, pp. 211–264). San Diego, CA: Academic Press.
Whittlesea, B.W.A. (1997b). The representation of general and particular knowledge. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 335–370). Hove, UK: Psychology Press.
Whittlesea, B.W.A., & Dorken, M.D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122*, 227–248.
Whittlesea, B.W.A., & Dorken, M.D. (1997). Implicit learning: Indirect, not unconscious. *Psychonomic Bulletin & Review, 4*, 63–67.
Whittlesea, B.W.A., & Williams, L.D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica, 98*, 141–165.
Wright, R.L., & Whittlesea, B.W.A. (1998). Implicit learning of complex structures: Active adaptation and selective processing in acquisition and application. *Memory & Cognition, 26*, 402–420.

CHAPTER FIVE

Knowledge representation and transfer in artificial grammar learning (AGL)

Martin Redington
University College London, UK

Nick Chater
University of Warwick, Coventry, UK

In this chapter we re-evaluate the implications that have been drawn from the artificial grammar learning transfer paradigm and review the evidence on whether the knowledge acquired in artificial grammar learning (AGL) is stored in terms of the original surface form or in an abstract, surface-independent representation.

INTRODUCTION

Research on AGL (Reber, 1967, 1993) has been seen as having potentially far-reaching implications for theories of cognition. These studies have appeared to provide evidence that people are able to learn "implicit" rules, that is, they can learn sets of rules without, in some sense, conscious awareness, and these rules can determine their behaviour. In consequence, AGL research has attracted a great deal of excitement and controversy.

A typical AGL study proceeds as follows: Participants are asked to memorise strings that (unknown to them) have been generated by a finite-state grammar, which dictates the permissible arrangements of letters. After training, participants are told that the first set of stimuli was rule-governed, and are asked to indicate which of a second set of strings obey or violate the rules. Half of these test strings are "grammatical" according to the rules that generated the first set of strings, and half are ungrammatical. In the standard AGL paradigm, with typical materials, participants will get about 65–75% of these classifications correct.

Perhaps the most obvious inference to draw is that people have, to an imperfect degree, learned the underlying rules that governed the strings. And strong claims about the implicit or unconscious character of these putative rules appear warranted, given that participants often declare that they have no idea whatever what rules governed the stimuli. Indeed, if pressed to make suggestions concerning this underlying structure, their suggestions, if any, tend to be partial and vague, and in any case insufficient to account for their performance.

It has, however, become increasingly evident that results of this kind are open to a wide range of interpretations. For example, it could be that people are simply memorising some of the original training items (Brooks & Vokey, 1991), or perhaps that they are learning fragments of these items, such as letter pairs or triples (e.g. Perruchet & Pacteau, 1990; Redington & Chater, 1996). People could be judging the new strings with which they are presented as "conforming to the rules" if they are similar either to specific old strings that they were presented with or if they contain similar fragments to those items. Straightforward analysis of typical AGL stimuli has shown that methods of this kind could give rise to high levels of performance in classifying new stimuli, and hence could explain human transfer performance. These exemplar- or fragment-based models would also explain immediately why people are able to say nothing about the rules governing the stimuli that they originally learned. They cannot report the rules, not because rule-knowledge is implicit but because their performance is not based on rule-learning at all.

There is, however, a version of the AGL paradigm that does not appear so easily open to alternative explanation by exemplar- and fragment-based accounts. It is the interpretation of these 'transfer" studies with which we are concerned in this chapter.

WHAT IS TRANSFER?

In the transfer paradigm, the surface form of the materials is changed between training and test. For example, the training strings MXVVM and VXMTM might become HJKKH and KJHLH, respectively. The new strings follow the same underlying rules as the original strings, but are expressed in terms of a completely new set of letters. Despite this change in surface form, participants are still able to distinguish test strings that follow the underlying rules from those that violate them at well above chance performance, and above the level of the performance of untrained control subjects.

Transfer has been demonstrated repeatedly (Altmann, Dienes, & Goode, 1995; Brooks & Vokey, 1991; Gomez & Schvaneveldt, 1994; Knowlton & Squire, 1996; Manza & Reber, unpublished manuscript; Mathews, Buss, Stanley, Blanchard-Fields, Cho, & Druhan, 1989; Whittlesea & Dorken,

1993; Shanks, Johnstone, & Staggs, 1997), and appears very robust. As well as transfer across letter sets, transfer has also been demonstrated across modalities (e.g. from sequences of letters to sequences of tones, and vice versa; see Altmann et al., 1995).

Transfer effects tend to be small and with such small effects careful control conditions are required (Redington & Chater, 1996). For example, one useful control is to ask participants who were not exposed to the original training stimuli to attempt to discriminate grammatical from ungrammatical stimuli. It might appear that such participants can only perform by chance guessing, and hence that such a condition is unnecessary but, in fact, such control conditions can yield performance that is consistently slightly above chance. This is possible because some aspects of the structure of the strings can be learned "on-line" as the participant inspects the test strings. After all, half of the strings are generated by the grammar and the other half are only slightly corruptions of grammatical items, and hence a good deal about the grammar can be learned as testing proceeds. Such controls are not always run in transfer studies and the possibility therefore remains that some apparently significant instances of transfer from the original learning items can be explained away as examples of learning during test. None the less, despite this and other methodological issues that can be raised in relation to some of the AGL transfer studies (Redington & Chater, 1996), the weight of evidence seems to strongly suggest that transfer effects are not merely artifacts—it appears that they genuinely reflect the application of some knowledge learned in memorising the training stimuli.

WHY IS TRANSFER IMPORTANT?

As we have indicated, transfer studies have assumed a particular significance in AGL research, in the light of the multiple interpretations available for conventional AGL experiments. We will focus here on two specific theoretical issues for which evidence from transfer has been viewed as important. Both of these claims concern the nature of the knowledge that people are acquiring when learning the training items.

Firstly, the transfer paradigm has been seen as evidence for the claim that participants in AGL studies acquire rule-based knowledge. That is, some theorists (e.g. Knowlton & Squire, 1996) have claimed that transfer could only be explained by assuming that participants learn rules that capture regularities in the items memorised during training. According to this viewpoint, transfer results rule out theories of AGL that propose that participants acquire knowledge of whole instances, or fragments of the training materials, represented in terms of the original surface form of those materials (e.g. Brooks & Vokey, 1991; Perruchet & Pacteau, 1990).

Secondly, transfer has been commonly interpreted as evidence for the claim that the knowledge acquired in AGL is represented not in terms of the original surface form of the training materials but, instead, in terms of an abstract "surface-independent" form (e.g. Mathews, 1990; Knowlton & Squire, 1996). For example, participants might genuinely acquire a representation of the finite-state grammar underlying the training items, with the arcs between the nodes of the grammar marked with abstract symbols that do not denote specific surface features but which can be bound to any surface form. The logic underlying the claim that knowledge is represented in a surface-independent form is that this knowledge can then be applied to stimuli presented in any surface form, as shown in the transfer paradigm.

Thus, the claim is that transfer effects demonstrate that the knowledge learned in encoding the training items is represented in terms of rules, and that it is encoded in a surface-independent format. These claims, and others that follow from them concerning conscious awareness, have given a substantial impetus to the study of transfer in AGL. But, we shall argue, neither of these putative implications of transfer is valid.

WHAT DOES TRANSFER REALLY SHOW?

We will first question the notion that transfer is evidence for rule-based knowledge, and secondly that it is evidence for the acquisition of surface-independent knowledge.

Transfer and rule-based knowledge

The key point here is that surface-independence and rule-based knowledge are orthogonal concepts. Within the AGL literature, discussion has concentrated on three distinct kinds of representation: (1) knowledge of whole exemplars (e.g. Brooks & Vokey, 1991); (2) knowledge of fragments of the training items (e.g. Perruchet & Pacteau, 1990); (3) or rule-based knowledge (e.g. Knowlton & Squire, 1996).

We argue that all three kinds of knowledge can, in principle, be tied to particular surface forms or, alternatively, can discard information about particular surface forms, resulting in a surface-independent representation.

Considering exemplar-based knowledge first, in Brooks and Vokey's (1991) account, participants' knowledge consists of representations of entire training items (such as the string MMTTVVX), with grammaticality judgements being based on the whole item similarity between the test items and the memorised training items. In Brooks and Vokey's account, the representations of these exemplars are expressed in terms of the original surface form, with transfer being accounted for by a process of "abstract analogy" between different surface forms.

In principle, one can conceive of an alternative exemplar-based mechanism, in which participants acquired representations of specific training exemplars but discarded, abstracted away from, or simply forgot, the particular surface features that were present during the training phase. Such a mechanism might acquire representations such as □□◇◇△△▽ (where □, ◇, △, and ▽ are symbols that discard the original surface form [letters] of the item, but retain distinctions between different letters).

If people possessed such a mechanism (and, for the sake of illustration, its contents were accessible to verbal report), we might receive verbal reports like "I can remember the item, but not the actual letters. I know that it had three different letters, each repeated twice, and then a fourth different letter".

Similarly, fragment knowledge, as discussed in the AGL literature, is generally held to be represented in terms of the surface form of the training materials, concerning the occurrence and potentially frequency of occurrence of given fragments (such as VXX) within the training items. However, a mechanism that acquired a similar kind of knowledge but discarded the particular surface form of the relevant features is equally possible, *a priori*. Such a system might learn that the fragment □◇◇ occurred 10 times in the training items, where □ and ◇ have the same meaning as above.

Within the AGL literature, the exact nature of rule-based knowledge is generally left rather vague. Two possible examples of rule-based knowledge are a representation of the finite-state grammar that actually generated the stimuli, and symbolic rules such as "All strings must begin with the letters VX or MX". Either kind of representation could be couched in terms of surface symbols, or in terms of abstract symbols. For example, a surface-independent symbolic rule might be expressed as something like "All strings must begin with one of two specific letters" or "The second letter is always the same".

Within the broad classes of theory defined by the categories of exemplars, fragments, and rules, and the orthogonal categorisation according to whether this knowledge is tied to surface form, or surface-independent, there is still room for a lot of variation and disagreement. For example, if people acquire surface-independent knowledge of exemplars, they might encode these in terms of the repetition patterns, discarding commonalities occurring across strings, so that the training items MMTTVVX and VTXXMV might be encoded as □□◇◇△△▽ and □◇△△▽□. Alternatively, the encoding might respect the identities of the original surface form across training items, so that the items above might be encoded as □□◇◇△△▽ and △◇▽▽□◇.

Additionally, there are accounts that do not fit quite so neatly into the division between exemplars, fragments, and rules, such as Dienes, Altmann, and Gao's (1999) connectionist model of transfer (although this could be argued to acquire knowledge that can be seen as similar in nature to fragments, or rules, according to one's theoretical perspective), as well as "hybrid" accounts that posit more than one form of knowledge (e.g. Knowlton & Squire, 1996; Meulemans & Van Der Linden, 1997).

Overall, though, our claim stands: Evidence that knowledge is surface-independent does not constrain that knowledge also to be rule-based. Exemplar- and fragment-based knowledge can also be represented in abstract terms, divorced from the original surface form of the training materials.

Transfer and surface-independent knowledge

We now turn to consider the implications of transfer for the level of abstraction of information that participants perform during training. In particular, we shall argue that although transfer is compatible with the hypothesis that knowledge is represented independent of surface form, it is not direct evidence for that hypothesis. Rather, there are many accounts in which participants' knowledge is held to be surface-based that can explain the phenomenon of transfer.

Before considering how surface-based knowledge might allow transfer, let us first consider how surface-independent knowledge might support transfer. This could happen in a number of ways.

In one possible account, both training and test items are encoded in a common surface-independent form, and thus the representations of their structure can be compared directly. In its purest form, this account requires that the mapping from surface to abstract form can be computed independently for each item, without possessing any knowledge about the other items. The example given above, where exemplars are encoded in terms of their repetition patterns falls into this category.

In another possible account, people initially acquire surface-independent knowledge and, in the course of testing, learn to map from the surface-independent form to the new surface form, or vice-versa, and thus exploit their knowledge for the purpose of the categorisation task. Note that the mapping between abstract and surface forms at test need not be consistent across items, or even within a particular item—in Roussel and Mathews' account of transfer (cited in Dienes & Berry, 1993), participants' knowledge consists of information concerning whether each letter is the same as, or different from, the preceding letter. All that is required is that the mapping is sufficient to provide some indication of the grammaticality of the test items.

Now consider the case where participants' knowledge is represented purely in terms of the original surface form. Although intuitively such knowledge would appear to be relatively useless for the purposes of transfer, a number of accounts of transfer have argued that participants' knowledge is surface-based.

Brooks and Vokey (1991) argued that knowledge of training exemplars could be applied to transfer test items by a process of "abstract analogy": The test item HHJJKKK is "similar" to the exemplar MMTTVVV, in that they share a common repetition pattern. While the underlying basis of comparison is abstract and unrelated to surface form, Brooks and Vokey are explicitly committed to the view that the knowledge acquired during training is instance-based, and closely tied to the original surface form of the training items. The abstraction that is required to support transfer takes place between surface forms, at test.

Similarly, Redington and Chater (1996) proposed an idealised model of transfer in which participants acquire knowledge of fragments of the training items, expressed in terms of the original surface form. In the transfer paradigm, test items are accepted as grammatical if there is a consistent mapping between old and new surface forms (computed on an item-by-item basis) such that the test item can constructed entirely from the fragments seen during training.

Finally, Dienes et al. (1999) proposed a connectionist account of transfer. The network is trained to produce responses that are specific to a particular surface form. These responses are predictions about which letter will occur next in the string—by comparing the predictions against the actual next letter, the network produces a measure of the extent to which the test item shares a common underlying structure with the training materials. In the context of transfer, the network is effectively "re-trained", but its architecture allows a new input–output mapping to use some of the same weights that were trained initially. With these "core" weights (whose values are frozen) already trained, the learning problem collapses into finding an appropriate mapping from the new input and output units to and from the core weights. This process can be viewed as mapping knowledge that is "bound" to one surface form to the new surface form.

In all three accounts, the knowledge acquired during training is bound to the original surface form of the training items. At test, some form of mapping between the old and new surface form allows this knowledge to be applied to the new materials. Whether, or to what extent, any of these three accounts can explain human performance in the AGL paradigm is open to question. However, they show that theories positing the acquisition of knowledge that is represented in terms of the original surface form are viable, and compatible with the basic fact of transfer of knowledge to a new surface form. This is especially true of the Dienes et al. (1999) and Redington and Chater (1996) work, where computational models were demonstrably able to perform the transfer task, fitting the details of the human with varying degrees of accuracy.

TABLE 5.1.
Actual and possible accounts of transfer in AGL, categorised by the kind of knowledge (rule, instance, or fragments) and whether that knowledge is held to be surface-independent or surface-based.

	Surface-independent	*Surface-based*
Rules	Reber (1969)	
Instances		Brooks & Vokey (1991)
Fragments		Redington & Chater (1996)

KNOWLEDGE REPRESENTATION IN AGL

Accepting that transfer is not, of itself, evidence for the acquisition of rule-based or surface-independent knowledge really opens up the field for possible theories of transfer, and AGL in general. If one also accepts that these concepts can be considered somewhat independently, then we can categorise accounts of knowledge representation in AGL into six categories, as shown in Table 5.1.

A note of explanation is required for our classification of the Dienes et al. (1999) connectionist model as "fragment-based", and exactly what we mean by this term. We regard "fragment-based knowledge" as shorthand for learning about simple local, distributional properties of the stimuli, such as bigram statistics—the frequency with which a particular pair of adjacent letters occurred in the training materials. "Localness is logically . . ." is logically independent from the learning of distributional properties but, in practice, most of these accounts propose that what is learnt is the distributional relationships between adjacent or nearby elements of the stimuli (e.g. letters in a string), which tallies with the observation that, with typical AGL training materials, people are more sensitive to violations of local dependencies between letters (e.g. Gomez & Schvaneveldt, 1994). This is the kind of information that the Dienes et al. (1999) connectionist network model acquires (see Chater & Conkey, 1993), and this kind of knowledge is also captured by models that acquire knowledge of fragment occurrence, as in the idealised models proposed by Redington and Chater (1996).

Obviously, the categories in Table 5.1 are somewhat coarse, and there is some scope for debate about their exact boundaries, or which category a particular account belongs in, but generally they reflect the broad theoretical divisions that exist in the literature. Most accounts of implicit learning can be placed into one of these categories. Where hybrid theories propose more than one distinct type of knowledge, those types can themselves each be classified into the categories above. For example, Knowlton and Squire (1996) and Meulemans and Van Der Linden (1997) both propose that that AGL standard and transfer phenomena can be accounted for

in terms of the acquisition of both surface-based fragment knowledge, and surface-independent rule-based knowledge.

A division within the literature that is not reflected within the categories above is on the issue of conscious awareness and accessibility of knowledge, which has been a major source of debate within the AGL literature.

We will not focus directly on the question of conscious awareness here because we believe that to a first approximation, how knowledge is represented can be considered independently from whether that representation, or aspects of it, are available to conscious awareness or not (although we concur with Dienes & Perner's 1996 suggestion that some representational forms might be more accessible or transparent to conscious awareness). Indeed, addressing the question of how knowledge about the training items is represented seems to be prior to addressing the question of whether such knowledge is conscious or not. Thus, if knowledge is believed to be represented in terms of rules of a particular kind then we can test, experimentally, whether people are aware of these rules; and if knowledge is represented in terms of fragments, then we can test whether they are aware of such fragments. But it seems to be inappropriate to attempt to address the question of conscious awareness before at least some clarity has been achieved concerning what kind of knowledge may or may not be consciously available.

For the purposes of this chapter, we will also leave aside the question of whether the knowledge acquired in AGL concerns exemplars, or fragments, or is rule-based. Instead we will focus below on the question: Is the knowledge acquired in AGL represented in terms of the original surface form, or in a surface-independent terms?

SURFACE-INDEPENDENT AND SURFACE-BASED REPRESENTATIONS

We have argued that the mere existence of transfer does not immediately imply that the knowledge learned in AGL training is necessarily surface-independent. This raises the question of how we can obtain evidence concerning the level of abstraction of AGL knowledge.

We have previously (Redington & Chater, 1996) discussed this question under the heading of the locus of abstraction in transfer—whether, in the transfer paradigm, there was abstraction away from the original surface form (resulting in a surface-independent representation), or whether (a slightly different kind of) abstraction took place between surface forms at test. Differentiating between these two possibilities is a specific instance of a problem that arises whenever a task requires abstraction of some kind, whether across training instances or away from a particular surface form.

From performance on the task itself, it is not always possible to determine logically whether participants acquired a representation in an appropriately abstract form during training, or whether the necessary abstraction occurred during testing, driven by the demands of the test procedure.

This does not mean that no evidence can be brought to bear. Redington and Chater (1996) proposed that there were three potential lines of evidence for surface-independent representations: (1) verbal report; (2) parsimony; and (3) evidence from indirect tests.

Considering evidence from verbal report first, if it were the case that participants in AGL studies clearly and unambiguously reported knowledge of surface-independent properties of the training stimuli, then this would provide some grounds for believing that participants' knowledge was represented in a surface-independent form. Unfortunately in AGL, as we have already noted, verbal reports are typically not very informative about the extent and nature of participants' knowledge (e.g. Reber & Allen, 1978), and so the evidence from verbal report favours neither a surface-based or surface-independent representation.

Turning to parsimony, all else being equal, it could be argued that accounts positing a surface-independent representation, which allows training and test items to be compared directly, provide a simpler explanation of transfer results than accounts that posit a process of abstraction between surface forms at test. But such an appeal to parsimony is problematic in the light of the consistent experimental findings that, given the same materials, classification performance in the standard (same letter-set) AGL paradigm is reliably higher than in the transfer case (e.g. Altmann et al., 1995; Manza & Reber, 1994)—it appears that participants do retain at least some information about surface form, and can take advantage of this for the classification task. For example, in Manza and Reber's (1994) account of the differential performance between the standard and transfer cases, participants acquire both surface-based knowledge and a subset of surface-independent knowledge. It seems that some surface-based representation will always be required to account for this difference.

Now, parsimony threatens to work against advocates of surface-independent knowledge: Regardless of whether knowledge is represented in surface-based or surface-independent terms, some process of "abstraction" operating either on the acquired representation, or the test item, or both, is required at test to allow the existing knowledge to be applied to the novel surface form. Given that surface-based knowledge has been shown to be sufficient to account for transfer performance, and that some surface-based knowledge is required to account for superior performance when the surface form of the materials is unchanged between training and test, accounts positing a single surface-based representation are arguably simpler than accounts positing multiple representations.

Redington and Chater's (1996) idealised computational models of AGL provide an existence proof for this kind of account, as does the connectionist model of AGL presented by Dienes et al. (1999).

In Redington and Chater's (1996) models, knowledge was represented as knowledge of the fragments (letter pairs and triples) that occurred in the training items, and this knowledge could be applied either directly to the test items (in the standard AGL paradigm) or by a simple process of finding correspondences between each test item and the training items (in the transfer paradigm). These models were able to capture the basic facts of transfer performance, and superior standard (non-transfer) performance, utilising only a single, surface-based representation.

Of course, parsimony might provide an *a priori* reason to favour one account over another, but empirical evidence provides a much more powerful argument. Redington and Chater (1996) suggested that indirect, or incidental, tests of performance, which avoided placing demands on participants that might lead to abstraction during the test phase, could provide evidence for an abstract representation, but failed to identify any converging evidence in the literature. This appears to leave the question of the level of abstraction of knowledge in AGL entirely unresolved.

EMPIRICAL EVIDENCE

There is, however, an early implicit learning study performed by Reber (1969) that does appear to provide strong evidence for the acquisition of surface-independent knowledge.[1] We first describe Reber's study, before considering why it provides a better test of the level of abstraction of AGL knowledge than later transfer studies.

In Reber's (1969) study, in contrast to later transfer studies, there was no test phase where participants classified test items as conforming to or violating the rules underlying the training items. Instead, performance in Reber's study was measured by the number of recall errors that participants made while memorising sets of training items. Each set (of three training items) was presented until the participant successfully recalled all the items on a single trial. The "test phase" of the study consisted of memorisation of a second set of training items. If participants were acquiring knowledge of the underlying structure of the materials during the first phase of the study, then this should help them to memorise items generated by the same

[1] Whittlesea and Wright (1997, pp. 190–191) similarly identify Reber's (1969) study as a crucial piece of evidence in favour of the acquisition of surface-independent knowledge, using a similar argument to that outlined here.

grammar in the second phase, providing a "memorisation advantage" for learning these new items.

Reber manipulated the relationship between the training items used in the first and second phases of the study: Whether or not they were generated by the same underlying grammar or not, and whether or not the items were expressed in terms of the same surface form. These factors were varied independently, to produce four experimental groups: same-grammar/same-letters, same-grammar/different-letters, different-grammar/same-letters, and different-grammar/different-letters.

As in previous studies (Miller, 1958; Reber, 1967), participants in the same-grammar group showed a memorisation advantage (made fewer errors when memorising the second set of items) over the different-grammar groups. The crucial result in Reber's (1969) study is that this advantage was maintained even when the surface form of the second set of items was different to that of the first set.

In the same-letter set case, this finding has been explained in terms of an encoding used for the first set of items being available to encode the second set of items, leading to more efficient memorisation (Miller, 1958; Reber, 1967). Reber's (1969) result indicates, as do the transfer results in the context of the classification task, that this encoding can be applied independently of surface form. However, Reber's study differs significantly from the grammaticality judgement studies in that participants in Reber's study knew nothing about the underlying structure of the training items throughout the experiment. Specifically, they did not know that any of the items were generated by a set of rules, or that the second set of items might bear any relationship to the first set.

Participants' naivety as to the letter set change is a critical difference. Accounts of transfer that do not posit a surface-independent representation (e.g. Brooks & Vokey, 1991; Redington & Chater, 1996) propose that a process of abstraction between surface forms takes place at test, and within these accounts this is typically seen as driven by the knowledge that there is a given relationship between training and test items. In the Dienes et al. (1999) model, for example, a set of "core weights" is set during training, but during the classification phase, the values of these core weights are frozen and used to constrain the training of a set of mapping weights (allowing the utilisation of the knowledge embodied in the core weights with the new surface form). In the context of this model, knowledge of the relationship between the training and test items can be seen as the signal to freeze the core weights and start the mapping process. Similarly, in Redington and Chater's (1996) idealised models, a process of mapping occurs at the test phase but there is no process for exploiting knowledge to a different surface form during the training (memorisation) phase. While these accounts are *a priori* compatible with the memorisation advantage in the same-letter set case, they cannot account for the Reber (1969) finding across different letter sets.

Surprisingly, although often cited, the Reber (1969) study has rarely been discussed in detail in the literature. While the reliability of the transfer effect itself has been questioned (Perruchet, 1994), and subsequently proved to be robust, no replications of the Reber (1969) effect have ever been reported.

We recently (Redington & Chater, unpublished data) performed two studies that replicate the underlying logic of the Reber (1969) study. In our first experiment, participants memorised sets of 18 strings in each phase. As in Reber's study, the grammar used to generate the strings and the letter set in which they were instantiated were varied between phases. Within each phase, the strings were presented in groups of three and participants had to correctly recall all three strings in a single trial before proceeding to the next group. The grammars that were used to generate the materials were not exactly the same as those used by Reber (1969) but were similar in nature and complexity (Figure 5.1), and had been constructed to minimise the sharing of repetition patterns between grammars: Of the 43 strings of length 8 or less that each grammar could generate, only 6 shared a common repetition pattern with the strings from the other grammar.

Participants in all four groups showed a statistically reliable improvement, both in terms of the number of presentations that they took to reach criterion for each group of three strings, within each phase of the study, and between the first and second phases. A reliable interaction indicated that the improvement over presentations (groups of three strings) was greater during the first phase of the study.

However, we found no evidence for the effect observer by Reber (1969). In Reber's study, the rank ordering of the performance of the experimental groups in the second phase of the study was as follows:

same-grammar/same-letters > same-grammar/different-letters >
different-grammar/different-letters > different-grammar/same-letters

In our study, the rank ordering of the groups was markedly different:

same-grammar/same-letters > different-grammar/different-letters >
same-grammar/different-letters > different-grammar/same-letters

The only reliable differences in memorisation advantage between the groups in our study were between the same-grammar/same-letters group and the other groups. Participants who studied items from the same grammar in both phases but instantiated in different letter sets, showed no memorisation advantage relative to the other groups.

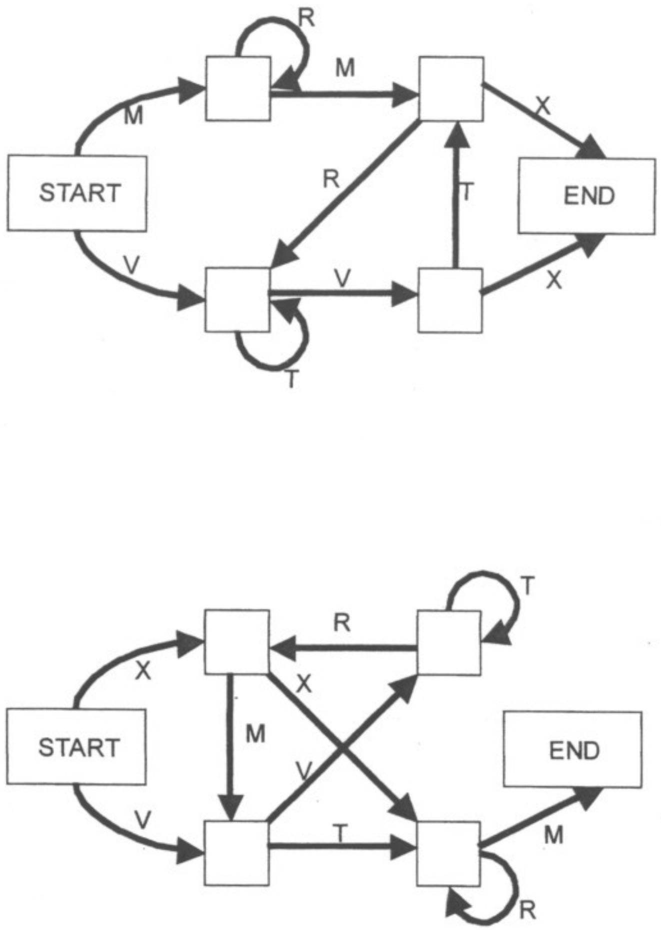

Figure 5.1. The two finite-state grammars in Redington and Chater's unpublished replication of Reber (1969). The strings are generated by beginning at the "start" state and following arrows through the diagram to the "end" state. Traversing an arrow corresponds to generating a symbol.

In a second study, we again presented participants with a memorisation task in two distinct phases, with the underlying grammar and surface form of the training strings varying across phases, as before. The procedure we used in this study was based on that of Miller (1958): Within each phase, the stimuli consisted of a list of nine strings—one set was generated by the grammar used by Miller (1958) and the other set was generated by a

grammar of similar complexity. Two different sets of nine strings were chosen from each grammar selected so as to be representative of the grammar, and so that their repetition patterns did not overlap with those of the strings selected from the other grammar. On each trial, all nine strings from the appropriate set were presented, in a random order, and the participants' task was to correctly recall as many of the items as possible. Each phase of the study consisted of ten trials, with the dependent variable being the number of correct responses on each trial.

As before, participants' performance improved reliably over trials and between the first and second phases of the study, with the improvement over trials being greatest during the first phase. The manipulations of grammar and surface form resulted in same rank ordering of the groups as our previous study:

same-grammar/same-letters > different-grammar/different-letters >
same-grammar/different-letters > different-grammar/same-letters

Once again, the differences between the same-grammar/same-letters group and the same-grammar/different-letters and different-grammar/same-letters groups were reliable, although the difference between the same-grammar/same-letters group and the different-grammar/different-letters group was not.

Hence, we failed to find any evidence in support of Reber's (1969) effect: There was no memorisation advantage for materials generated using the same grammar when the surface form was altered, relative to materials generated by a different grammar.

Of course, it is possible that the use of different stimuli could explain our failure to replicate Reber's (1969) findings, although other implicit learning effects, such as classification in the standard and transfer AGL tasks, are usually robust in the face of the kind of variation this introduced.

If our results are valid, it seems that the reliability of Reber's (1969) finding is at least open to question and that under "normal" AGL learning conditions, with adult participants, given straightforward memorisation instructions and a moderately complex set of rules governing the materials, the empirical evidence appears to be consistent with the acquisition of a representation that is tied to the original surface form, and does not favour a surface-independent representation.

This conclusion is further reinforced by a more recent transfer study reported by Whittlesea and Wright (1997, Experiment 4). As in a normal transfer study, participants were required to memorise a set of training strings and then, after being informed of the rule-governed nature of the materials, and that the training and test materials shared a common set of underlying rules, they performed a classification test on items that were instantiated either in a different letter set, or as sequences of coloured patches (with each colour corresponding to one of the letters in the original stimuli). Prior to being informed about the nature of the materials and performing the classification task, participants were asked to judge each of a set of test items, instantiated in the same form as the items on the classification task they were about to perform, as being pleasant or dull.

Whittlesea and Wright (1997) chose the pleasantness judgement because of its use as a measure of the mere exposure effect (Zajonc, 1968): Stimuli to which people have been previously exposed are judged as more pleasant than novel stimuli. With standard (non-transfer) AGL materials, Gordon and Holyoak (1983) found that participants were likely to rate grammatical test items as more pleasant than non-grammatical test items.

Just as in the memorisation task used by Reber (1969), the pleasantness judgement task does not require participants to be aware, or informed, that the stimuli are rule-governed, or of the relationship between the training and test stimuli. This decreases the possibility that abstraction to a new surface form is driven by the demands of, and occurs during, the test phase, allowing a reasonable test of the hypothesis that participants' knowledge is represented in surface-independent terms. Participants acquiring surface-independent knowledge should show effects of grammaticality on both the pleasantness and classification tasks under transfer conditions. On the other hand, if people possess only surface-based knowledge then they can use this knowledge to obtain better-than-chance levels of performance in the transfer classification task, by using abstraction at test, but there should be no transfer effects on pleasantness judgements.

The latter pattern of results was observed. Whittlesea and Wright (1997) reported that their participants were able to classify items as grammatical or non-grammatical at above-chance levels when the surface form of the test items was instantiated in a different letter set to that used in training, or as a series of colour patches. However, on the pleasantness task, participants who were presented with sequences of coloured patches at test showed no sensitivity to the grammaticality of the test items, suggesting that they were not exploiting surface-independent knowledge when performing the pleasantness task.

Participants who were presented with strings of letters (instantiated in a different letter set) were reliably more likely to classify grammatical items as more pleasant than non-grammatical items. Whittlesea and Wright (1997) suggest that this does provide evidence that memory is "directly sensitive to the deep structure [surface-independent properties] of the environment" (p. 194). However, it should be noted that the effect is rather small. Only 53% of pleasant or dull classifications corresponded with the items' grammaticality or non-grammaticality (where 50% would be expected by chance). In contrast, their grammaticality judgement score was 60.1%. Additionally, the comparison of participants' performance was against chance rather than, for example, against a control group, leaving open the possibility that at least some of the observed effect was due to participants' pre-existing biases, rather than exposure to the training materials (see Redington & Chater, 1996, for a discussion of this and other possible biases in AGL). Whittlesea and Wright also tested an untrained group, whose pleasantness task performance was also above chance, although not reliably so, but did not compare the performance of the experimental groups against this control.

To summarise, although Reber's (1969) results (which showed a memorisation advantage for participants who had been exposed to stimuli with the same underlying structure but with a different surface form) suggest that people do acquire surface-independent knowledge, we were unable to replicate this finding. Our results indicated that once the surface form of the materials is changed, any memorisation advantage for previous exposure to the grammar disappears. Whittlesea and Wright's (1997) results, using a similarly indirect method to detect the influence of surface-independent knowledge, also show no effect of such knowledge in the case where the surface form of the stimuli was changed from letters to colours, and only a very small (and potentially questionable) effect for letter–letter transfer. All in all, there appears to be little empirical evidence for the acquisition of surface-independent knowledge in adults, in conditions typical of the artificial grammar learning studies: Straightforward memorisation of moderately complex sequential materials.

EVIDENCE FOR THE ACQUISITION OF SURFACE-INDEPENDENT KNOWLEDGE

Although adults might not acquire surface-independent knowledge in "normal" implicit learning conditions, Whittlesea and Dorken (1993) provide some evidence that they can acquire surface-independent representations in some circumstances. Another intriguing line of evidence (Gomez & Gerken, 1999; Marcus Vijayan, Bandi Rao, & Vishton, 1999) suggests that young infants could acquire surface-independent representations, even if adults do not.

In the Whittlesea and Dorken (1993, Experiment 5) study, participants were trained and tested in the standard and transfer AGL paradigms, with a learning phase consisting of exposure to the training materials, making grammaticality judgements about novel strings during the test phase. The important experimental manipulation was in the nature of the training, which was intended to influence the way that participants processed the training items. One group of participants (the incidental repetition group) was told that the training items were distractors in a digit repetition task—they had to repeat a series of letters (i.e. a training item) between presentation and recall of a three-digit number. In the incidental analysis condition, participants' training consisted of a "repetition-detection" task during training: A letter of the training item would be underlined and the task was to indicate whether that letter occurred at any other location in the test item, as quickly as possible. A third group (the memorise group), performed a straightforward memorisation task.

Whittlesea and Dorken found that classification (grammaticality judgement) performance in the standard (non-transfer) conditions was lowest for the incidental repetition and incidental analysis groups, with the memorisation group performing reliably better. In transfer test conditions, the performance of the incidental analysis group, whose training task focused attention on the repetition patterns of the training items, reliably exceeded that of the memorisation task.

The superior transfer performance of the group trained to detect repetition patterns in the training items suggests that, under these training conditions, participants might have acquired surface-independent knowledge during training. It seems unlikely that such a training regime, if it resulted in knowledge that was completely tied to surface form, could lead to improved transfer performance at test. However, the training regime used here is very different to straightforward memorisation: Given that the task directed participants to attend to the repetition patterns of the training items, it is hardly surprising that they acquired some knowledge of these patterns—the implications of this finding for artificial grammar learning occurring under "normal" conditions are limited.

However, recent studies reported by Gomez and Gerken (1999) and Marcus et al. (1999) suggest that even under "passive" learning conditions, human infants can acquire knowledge of surface-independent structure.

The Gomez and Gerken (1999) and Marcus et al. (1999) studies both utilised the Head Turn Preference Procedure (Kemler-Nelson, Jusczyk, Mandel, Turk, & Gerken, 1995). Infants sit on a parent's lap in a three-sided booth. Stimuli are played through speakers on either side of the booth, and the infants' tendency to attend to one side or the other, as indicated by head-turning or looking times, is used as a measure of their ability to distinguish between the stimuli.

This procedure allows an almost exact analogue of the adult artificial grammar learning studies to be performed with infant participants. Saffran, Aslin, and Newport (1996), reported a study on the acquisition of word segmentation, where 8-month-old infants were exposed to a continuous speech stream of four trisyllabic nonsense words presented in random order. Using the head-turn procedure, Saffran et al. showed that the infants could subsequently distinguish between words and part-words or non-words, as indicated by a novelty preference in their looking behaviour when the test stimuli were played.

Gomez and Gerken (1999) and Marcus et al. (1999) performed very similar studies to Saffran's, except that in their studies the surface form of the training materials was switched between the training and test phases of the study. Nevertheless, infants (7-month-olds in the Marcus et al. study, and 11-month-olds in the Gomez & Gerken study) showed some sensitivity to the underlying structure of the training materials. In the Marcus et al. study, infants showed a novelty preference, attending more to stimuli with the same underlying structure as the training materials. In the Gomez and Gerken study, infants preferred to listen to items that shared a repetition pattern with the training materials. The actual direction of the effects is unimportant to the argument presented here—in each case, infants were able to distinguish between items that shared the same underlying structure as the training items and those test items that did not.

Why is this result evidence for the acquisition of a surface-independent representation, when adult studies of transfer are not? Given the age of the infants, no meaningful instructions as to the rule-governed nature of the stimuli and the relationship between the training and test materials could be given. The infant participants were necessarily completely naive, providing an indirect test of the knowledge that they acquired during training.

One caveat to this conclusion is that, in the Marcus et al. (1999) studies, the stimuli were generated from two grammars that generated three-syllable words, all of which had the repetition structure ABB (in one grammar), or ABA (in the other grammar), and the 16 stimuli were each repeated three times in the course of a 2-minute training session. With such a simple structure, in adults at least, it would not be surprising if participants acquired explicit (in the sense of being readily verbalisable) knowledge of this repetition pattern, and exploited this in the test phase, and it seems plausible that a similar process of explicit learning (at some level) could take place in infants. However, the Gomez and Gerken (1999) materials were much more akin to the relatively complex materials used in adult studies, providing a more convincing case for the kind of implicit processing that is seen in adults (although in the case of infants this also seems to encompass surface-independent knowledge).

DISCUSSION

We have seen that the literature of AGL and related paradigms have not yet provided decisive evidence concerning the level of abstraction of information learned during training. Although the existence of transfer in AGL studies has sometimes been interpreted as unequivocal evidence for surface-independent representations, an alternative hypothesis remains open: That "abstraction" away from the surface representation occurs only at test, triggered by the knowledge that there is common structure between the training and test items. Studies in which participants are ignorant of the relationship between training and test items provide a means of distinguishing between these rival hypotheses. Reber (1969) described such a study, which appeared to show evidence for transfer, in a memorisation paradigm where people were unaware that the strings that they were asked to learn were governed by rules, let alone that these rules were in some cases analogous to rules governing previously learned items (the "transfer" condition, in Reber's study). But two attempts to replicate this result led to the opposite conclusion—that there is no evidence for transfer, and hence no evidence for surface-independent learning. Using a different design, Whittlesea and Wright (1997) provided evidence consistent with these results, finding no evidence of transfer for a pleasantness rating task where participants were ignorant of the relationship between training and test items.

However, it appears plausible that when the training task requires participants to attend to the repetition structure of the training materials, as in Whittlesea and Dorken's (1993) study, people will acquire a surface-independent representation. The evidence provided by Gomez and Gerken (1999) and Marcus et al. (1999) suggests that young infants might acquire surface-independent representations, even under passive learning conditions, with speech-like materials.

The debate concerning levels of abstraction in AGL can be viewed as an aspect of a much broader theoretical debate in cognitive science: to what extent is human learning "lazy" or "eager"? (Aha, 1997; Hahn & Chater, 1997, 1998.)

Lazy learning involves storing input material in a relatively unprocessed form; the cognitive work required to transfer this knowledge to some new context (e.g. generalising past experiences to a new situation) is applied only when this work needs to be done. This style of learning is "lazy" because cognitive work is done only when strictly necessary—otherwise, the learning items are simply stored. In cognitive science, lazy learning is exemplified by exemplar models of categorisation (e.g. Medin & Schaffer, 1978; Nosofsky, 1986), memory (Hintzmann, 1986), case-based reasoning (e.g. Kolodner, 1993), and analogy-based models of reading (Glushko, 1979) and morphological processing (Nakisa & Hahn, 1996).

By contrast, eager learning involves actively attempting to extract regularities from new items, as they are encountered. The model of the regularities that has been extracted can then straightforwardly be applied to new items, as they are encountered. Eager learning methods vary between methods that involve the attempt to seek symbolic rules with which to model the incoming data (Lavrac & Dzeroski, 1993; Thagard, 1988), and those that attempt to fit incoming data to some kind of probabilistic model (e.g. a mixture of Gaussians, as in the decision-bound model of categorisation, Ashby, 1992).

It seems likely that the cognitive system uses some mixture of lazy and eager learning processes. For example, given the creativity of language use, it seems at best highly unlikely that the syntactic structure of a language can be learned using purely lazy learning. Instead, it seems possible to learn syntactic rules, which can then be used in ways that can be different from the contexts in which they were learned. But conversely, there is considerable evidence that the surface properties of learned items have a substantial effect on later memory and processing, in studies of categorisation, memory, and reasoning (e.g., Gick & Holyoak, 1980; Hintzmann, 1986; Posner & Keele, 1970).

To the extent that both lazy and eager processes might be operating in the cognitive system, it also seems plausible that both types of processing can be engaged in AGL studies. It might be that the abstractness and unfamiliarity of AGL stimuli, and the regularities defined over them, tends to block the application of eager learning methods. It might be that by using more natural, speech-like materials, as in the infant studies reported earlier, the cognitive system might be able to find abstract regularities during training. Alternatively it could be that learning styles shift during development, so that in the case of repetition structure, infant sensitivity to repetition structure is lost at some point during development.

The difference between the types of material, and experimental task, in which lazy or eager styles of cognitive processing are favoured, is a potentially important direction for future research, and AGL could be a valuable tool for investigating these questions.

ACKNOWLEDGEMENTS

Thanks to David Shanks and Theresa Johnstone for informative discussions on implicit, exemplar, and rule-based learning, and Pierre Perruchet for his comments on the manuscript.

This research was partly supported by ESRC research grant R000236214. Some of the work was performed while the authors were members of the Department of Experimental Psychology, University of Oxford.

REFERENCES

Aha, D.W. (Ed.) (1997). *Lazy learning*. Dordrecht, The Netherlands: Kluwer.

Altmann, G.T.M., Dienes, Z., & Goode, A. (1995). On the modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 899–912.

Ashby, F.G. (1992). Multidimensional models of categorization. In F.G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449–483). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Brooks, L.R., & Vokey, J.R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 316–323.

Chater, N., & Conkey, P. (1993). Sequence processing with recurrent neural networks. In M. Oaksford & G.D.A. Brown (Eds.), *Neurodynamics and psychology*. London: Academic Press.

Dienes, Z., Altmann, G.T.M., & Gao, S.J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science, 23*, 53–82.

Dienes, Z., & Berry, D. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Dienes, Z., & Perner, J. (1996) Implicit knowledge in people and connectionist networks. In G. Underwood (Ed.), *Implicit cognition* (pp. 227–256). Oxford: Oxford University Press.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.

Glushko, R.J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 674–691.

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., & Schvaneveldt, R.W. (1994). What is learned from artificial grammars? Transfer tests of simple associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 396–410.

Gordon, P.C., & Holyoak, K.J. (1983). Implicit learning and the generalization of the "mere exposure" effect. *Journal of Personality and Social Psychology, 45*, 492–500.

Hahn, U., & Chater, N. (1997). Concepts and similarity. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 43–92). Hove, UK: Psychology Press.

Hahn, U., & Chater, N. (1998). Understanding similarity: A joint project for psychology, case-based reasoning, and law. *Artificial Intelligence Review, 12*, 393–427.

Hintzmann, D.L. (1986). "Schema abstraction" in a multiple trace memory model. *Psychological Review, 93*, 411–428.

Kemler-Nelson, D., Jusczyk, P.W., Mandel, D.R., Turk, A., & Gerken, L.A. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behaviour and Development, 18*, 111–116.

Knowlton, B.J., & Squire, L.R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 169–181.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.

Lavrac, N., & Dzeroski, S. (1993). *Inductive logic programming: Techniques and applications*. New York: Ellis Horwood.

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

Mathews, R.C. (1990). Abstractness of implicit grammar knowledge: Comments on Perruchet & Pacteau's analysis of synthetic grammar learning. *Journal of Experimental Psychology: General, 119*, 412–416.

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.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207–238.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1007–1028.

Miller, G.A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology, 56*, 485–491.

Nakisa, R., & Hahn, U. (1996). Where defaults don't help: The case of the German plural system. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 177–182). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Nosofsky, R.M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 15*, 39–57.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 223–228.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119*, 264–275.

Posner, M., & Keele, S. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304–308.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 5*, 855–863.

Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81*, 115–119.

Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Reber, A.S., & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functional approach. *Cognition, 6*, 189–221.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General, 125*, 123–138.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical cues in language acquisition: Word segmentation by infants. In G.W. Cottrell (Ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 376–380). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Shanks, D.R., Johnstone, T., & Staggs, L. (1997). Abstraction processes in artificial grammar learning. *Quarterly Journal of Experimental Psychology, 50A*, 216–252.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.

Whittlesea, B.W., & Dorken, M.D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122*, 227–248.

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.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs, 9*(2. Pt. 2).