

Artificial grammar learning
Zoltán Dienes, University of Sussex
To appear in Oxford Companion to Consciousness. Oxford University Press.

A grammar is a way of specifying how elements should be ordered. One sort of grammar is a *finite state grammar*, as illustrated in Figure 1. It contains a finite number of states, represented by circles, connected by arrows showing allowable transitions. A grammatical sequence or ‘string’ can be produced by entering the network and changing states until the network is exited. So, for example, according to the grammar shown in the figure, XMMXM is grammatical.

In the 1960s Arthur Reber asked people to look at strings of letters generated by a finite state grammar (which was *artificial* in that it did not correspond to any natural language structure). Only then he informed his participants that the order of letters obeyed a complex set of rules. He showed people a new set of letter sequences, half of which obeyed the rules and half of which did not, and asked people to classify the sequences. He found people could classify with about 60-80% accuracy even though they found it very difficult to report knowledge of the rules of the grammar. Reber realised that such artificial grammar learning could be used as a method for investigating what he termed *implicit learning*, i.e. the acquisition of implicit (unconscious) knowledge (see Leaning, implicit vs explicit, this volume). In general, artificial grammar learning is investigated by any experimental method involving learning the structure of sequences of elements and would include, for example, the Serial Reaction Time task (see Sequence learning, this volume). However, we will deal with studies closely modelled on Reber’s original, but parallel conclusions can be drawn from the other paradigms. Research has addressed both what people learn and whether or not such knowledge is unconscious. We consider each in turn.

Initially, inspired by work in linguistics at the time, Reber believed people learned the syntactic structure of the grammar as such. As evidence, Reber showed that people asked to look at strings of letters generated by a grammar could later classify strings of completely different letters generated by the same grammar (e.g. if trained on strings using the letters M, T, V, R and X, they could be tested on strings made of the letters B, C, F, G and L). That is, there was knowledge that could be used when the grammar was the same but its concrete embodiment different. The first person to challenge the view that knowledge was of the grammar per se was Lee Brooks in 1978, who showed how people could memorise whole strings and classify by *analogy* to these remembered exemplars (e.g. MTTVT is very similar to MTTXV). People need not have learned any general rules just specific exemplars. Even when letter sets have changed, a test string MTTVT could be seen as abstractly similar to a training exemplar BFFCB, and correctly classified above chance without people inducing the rules of the grammar. According to this view much of human implicit or intuitive learning is based on retrieving relevant specific past cases (‘case based reasoning’), perhaps together with associated irrelevant aspects of context. Another view is that people learn chunks, for example of two successive letters (*bigrams* e.g. ‘MT’) or three successive letters (*trigrams*, e.g. ‘TTV’), or, in general, *n* successive letters (*n-grams*). There is considerable evidence that people do indeed learn which chunks are allowable in a range of implicit learning paradigms (see Leaning, implicit vs explicit, this volume). Both the above views (analogy to specific exemplars and chunking) are in principle consistent with the knowledge being conscious or unconscious. A third

view is that people's knowledge is embedded in the weights of a neural network and can thus fall along a continuum of concreteness or abstraction, depending on the details of the network. One neural network commonly used in the literature for simulating artificial grammar learning results is Elman's *simple recurrent network*, which can learn chunks as well as long distance dependencies. The view that the knowledge is embedded in weights fits naturally with the finding that the knowledge is hard to report verbally.

We turn now to research concerning the conscious status of the knowledge acquired in artificial grammar learning. The initial evidence for unconscious knowledge was that people's freely stated rules could not account for the level of classification performance achieved. But critics have been unhappy with free report as an indicator of unconscious knowledge. Conscious knowledge may not be freely reported merely because the person is not completely sure of it, has momentarily forgotten it, or does not believe it is the sort of knowledge the experimenter is looking for. One solution is to conjecture what properties of strings people have learned and give people forced choice tests concerning those properties. The use of forced choice tests ensures people do not hold back from using any knowledge they may have. For example, if it is conjectured that people have learned simply bigrams (e.g. that 'MT' can occur in grammatical strings) then people can be asked to discriminate legal from illegal bigrams. Indeed, people are good at discriminating legal from illegal bigrams, which has been taken to indicate that people's knowledge is conscious. The problem is that both conscious and unconscious knowledge would in principle allow such discrimination.

Another method for determining the conscious status of knowledge states is to ask the participant to report or discriminate not states of the world (e.g. 'can this bigram occur in the grammar?') but the relevant mental states involved in classification (see Objective vs subjective measures, this volume). Unconscious knowledge is knowledge a person is not aware of. Thus, we need to determine whether or not the person knows that they know in order to determine if the knowing is conscious. For example, confidence ratings can be elicited after each classification decision. According to the *guessing criterion*, unconscious knowledge is shown when the participant says they are purely guessing but they are in fact performing above baseline. According to the *zero correlation criterion*, knowledge is unconscious if the person cannot discriminate between when they are guessing and when they have knowledge, i.e. there is no relationship between confidence and accuracy. Both criteria have indicated unconscious knowledge, though a typical pattern is for there to be evidence of both conscious and unconscious knowledge.

The guessing criterion has been criticised because of the bias problem: When people say they are guessing, they might *think* that they are not. (Note that the existence of an adjustable bias for *thinking* one is guessing vs knowing is not in itself a problem for the guessing criterion.) A response to the objection has been to indicate evidence that people's reports of whether or not they are guessing distinguish knowledge types that differ in ways predicted by a theory of consciousness (e.g. resilience to a secondary task). The zero correlation criterion is less susceptible to the bias problem.

The guessing and zero correlation criteria measure the conscious status of *judgement knowledge*: i.e. knowing that this string is grammatical. That leaves open the question

of whether the person's *structural knowledge* (knowledge of the structure of the training strings) is conscious or unconscious. To address the latter, the experimenter can ask people after each classification decision whether they based their answer on random responding or intuition (unconscious structural knowledge) or rules or memory (conscious structural knowledge). This subjective method indicates people typically use both conscious and unconscious structural knowledge.

People trained on two grammars in different contexts can choose which of the grammars to use in the classification phase (when the content of their structural knowledge is sufficient for discrimination between the grammars) (see Memory, process dissociation procedure, this volume). That is, like bilinguals, people can choose which language to use; in that sense, grammatical knowledge is not applied automatically. Further, people trained on one grammar in one context do not apply it in a test phase in a different context unless told of the connection between the contexts. Such intentional control of the use of the knowledge often co-exists with lack of awareness of what the knowledge is by the guessing criterion. That is, a person does not need to be aware of controlling their knowledge nor of what the knowledge is in order to control it.

A common argument for there being unconscious knowledge learned in artificial grammar learning is that other primates and human babies as young as two months can learn statistical structures in sequences. The assumption made in this argument is that such creatures do not consciously test hypotheses nor do they have fully fledged episodic memory. Further, people with amnesia caused by damage to the temporal lobes learn artificial grammars almost as well as normal adults. But none of these facts entail that the corresponding learning mechanism in normal human adults produces unconscious knowledge. Moreover, the mere fact that a person has impaired episodic memory does not entail they do not use conscious knowledge, either judgment or structural (e.g. rules). However, these studies on different populations can be very informative about the basis of implicit learning in adults, when its implicit nature is established by other means.

Words: 1512

Further reading.

For an overview of Reber's work and ideas:

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

On the content of what is learned in artificial grammar learning:

Pothos, E. M. (2007). Theories of artificial grammar learning. Psychological Bulletin, 133, 227-244.

Subjective measures and judgment and structural knowledge:

Dienes, Z. (2007). Subjective measures of unconscious knowledge. In R. Banerjee & B. Chakrabarti (Eds), Models of Brain and Mind : Physical, Computational and Psychological Approaches. Elsevier.

Arguments against the evidence for unconscious knowledge:

Shanks, D. R. (2005). Implicit learning. In K. Lamberts and R. Goldstone, Handbook of Cognition (pp. 202-220) . London: Sage.

Artificial grammar learning in infants:

Gómez, R. L. (2006). Dynamically guided learning. In Y. Munakata & M. Johnson (Eds.) Attention & Performance XXI: Processes of change in brain and cognitive development (pp 87-110). Oxford University Press.

Figure 1. An example of a finite state grammar

