What Is Really Learned in Artificial Grammar Learning?
Implicit Intention for Learning in the Selective Attention Process

Daisuke Tanaka (dtanaka@ristex.jst.go.jp)
Center for Research on Brain-Science & Society, Research Institute of Science and Technology for Society,
Japan Science and Technology Agency, 1-1-2 Otemachi, Chiyoda-ku, Tokyo 100-0004, Japan

Sachiko Kiyokawa (kiyo@p.u-tokyo.ac.jp)
Graduate School of Education, University of Tokyo,
Hongo 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan

Ayumi Yamada (ayumi.yamada@gakushuin.ac.jp)
Department of Psychology, Faculty of Letters, Gakushuin University,
1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan

Kazuo Shigemasu (kshige@bayes.c.u-tokyo.ac.jp)
Department of Cognitive and Behavioral Science, Graduate School of Arts and Sciences
University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

Abstract

This study aimed to investigate what is really learned in artificial grammar (AG) learning and how selective attention influences AG learning. In the learning phase, depending on the aspects of GLOCAL strings—chains of compound letters that can allow two different AGs— participants were presented with two different series of letter strings. Their selective attention was manipulated by the experimental instructions. The results of the grammaticality judgment task following the learning phase revealed three findings. First, in accordance with the literature, selective attention was required for AG learning. Second, with regard to grammatical judgment, there was no difference between novel and presented strings, suggesting that grammatical judgment is based not on the representation of the presented strings themselves but on the AG extracted from them. Finally, there is evidence that selective attention works positively for the inhibition of the stimuli.

Keywords: artificial grammar learning; selective attention; implicit learning; global/local.

Background

The environment surrounding us contains almost infinite information for growth. However, organisms can identify only specific covariations as useful information for their adaptation. What is particularly interesting is that human beings, who tend to be interpreted as subjects controlled by their consciousness, can also learn implicitly. This ability is known as implicit learning (Reber, 1989). Since Reber published a series of pioneering researches using the artificial grammar (AG) learning procedure (e.g., Reber, 1967), many researchers have studied implicit learning, using several types of procedures, for example, the serial reaction time (SRT) and the AG learning procedures (Buchner & Wippich, 1998; Reber, 1989; Seger, 1994; Shanks & St. John, 1994, for reviews).
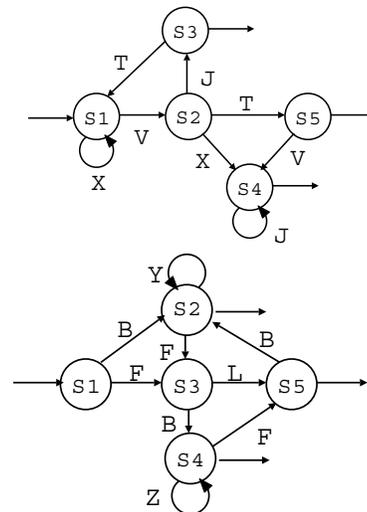

Figure 1: AGs used in this experiment.
Top panel: Grammar 1; Bottom panel: Grammar 2.

In the learning phase of the AG learning procedure, participants are exposed to a series of letter strings that follow complex rules—defined as the finite-state Markovian rule system (e.g., Figure 1). For instance, "XVJTVX" is presented as a grammatical string in the learning phase (see Grammar 1, the top panel of Figure 1). In the test phase following the learning phase, the participants are asked to select grammatical strings from novel strings, some of which follow AG rules and some violate them. For instance, "XTTVJJ" is considered as nongrammatical since the existence of the second letter "T" violates Grammar 1. The results reveal that participants can correctly select novel grammatical strings above chance level, even though they

cannot satisfactorily answer questions related to the selection criteria. Since they are unable to access the selection criteria explicitly, it can be speculated that they acquire the necessary knowledge implicitly.

One of the unique features of the AG learning procedure is the abstraction of rules extracted from the presented stimuli in the learning phase. In the SRT procedure, participants' responses to repeated patterns quicken regardless of their awareness of learning. This phenomenon can be interpreted as evidence that they can only store repeated patterns implicitly. In the AG learning procedure, however, participants in the test phase are able to distinguish grammatical strings from nongrammatical ones, even if they have never seen them before. It is considered that in the learning phase, participants can implicitly learn AG extracted from the presented strings, and they can subsequently respond based on the AG learned in the test phase.

What Is Really Learned through AG Learning

An alternative explanation about AG learning interprets participants' test phase performance as behavior based on the representation of presented strings in memory in the learning phase (e.g., Vokey & Brooks, 1992). This is contrary to the view that AG learning differs from the SRT procedure with regard to the extraction of rules from a series of presented strings.

This explanation leads to the following assumption: If participants' behavior is based not on the knowledge of AG but on the representation of presented individual strings, their responses to presented grammatical strings differ from those to novel grammatical strings. More specifically, it seems that in grammatical judgment, the rates of correct responses for presented letter strings are higher than those for novel grammatical strings. No difference is predicted between participants' performances with regard to the two types of strings if they can extract AG and use it in their grammatical judgment.

Relationship between Implicit Learning and Selective Attention

Jiang and Leung (2005) using visual search task as the implicit learning about a sequential pattern revealed that an unattended repeated pattern was accumulated as the consequence of latent learning, though this is not manifested in participants' behavior. It is very difficult to distinguish implicit learning from latent learning strictly with respect to the cognitive process. However, selective attention is not responsible for the simple storage of environmental information and may therefore execute the selection of information that is needed to behave appropriately, thereby making inputted information inhibit actively.

Tanaka et al. (2006) investigated the role of selective attention in AG learning using GLOCAL letter strings

(Figure 2). A critical feature of this stimulus is that while a GLOCAL string can be read as one string by using global letters (NVJTVJ in Figure 2), it can also be read as a string using local letters (BYYFLB in Figure 2). In other words, GLOCAL strings are chains of compound letters (Navon, 1977). A compound letter represents one large letter (i.e., a global letter) composed of a set of small letters (i.e., local letters). It is well known that when different letters are represented in local and global aspects, there exists interference between the two aspects during the preattention process: the irrelevant as well as relevant aspects cannot inevitably stop processing even if they are not attended to. Further, the direction of interference is asymmetrical: the effect of the global to local aspect is larger than that of the opposite direction. However, when the presentation time of stimuli is sufficiently long for recognition, bidirectional interference occurred (Hibi, Takeda, & Yagi, 2002). Tanaka et al. (2006) found that participants could learn only one AG from the attended aspect of the GLOCAL strings, although two aspects of GLOCAL strings provided enough information to learn two different AGs. They concluded that selective attention plays a critical role in AG learning.

```
B        B   Y       Y              Y   FFFFFFFFF   L       L             B
B B      B   Y     Y                Y   F           L       L             B
B  B     B   Y   Y                  Y   F            L     L              B
B   B    B   Y  Y                   Y   F             L   L               B
B    B   B   Y Y    Y               Y   F              L L    B           B
B      B B    YY      Y       Y      Y   F             LL    B            B
B        B    Y          YYYYY       F                 L       BBBBB
```
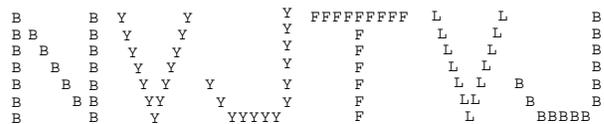
Figure 2: An example of GLOCAL strings.

In the research on the mere exposure effect that has the same paradigm with AG learning as implicit learning—except that it uses preferential judgments as the measurement of the test phase—Newell and Bright (2003) revealed that the structural mere exposure effect disappeared when the presentation time was brief, whereas the classical mere exposure effect was intact. In other words, brief presentation time interfered with the generalization of grammatical information from presented strings. Some studies, including Newell and Bright (2003), have shown that the process of generalization from presented strings differs from the process of storage for presented strings themselves (e.g., Kramer & Parkinson, 2005; Zizak & Reber, 2004). The result of Tanaka et al. (2006) provided two possible interpretations. First, selective attention affects the extraction of rules from the presented strings. It is therefore predicted that participants can make the representation in memory of presented strings of the unattended aspect of GLOCAL strings even if they cannot extract AG from them, because during the preattention process, the unattended aspect of GLOCAL strings should inevitably function as the feature of compound letters. Second, in AG learning, participants inhibit the unattended information positively rather than process it automatically; therefore, it is predicted that participants cannot even make the representation in

memory of presented strings of the unattended aspect of GLOCAL strings.

Although Tanaka et al. (2006) revealed the importance of selective attention in the AG learning procedure, their results could not discriminate between these two possibilities—failure of the extraction of AG from the unattended aspect and inhibition of the input from the unattended aspect as the effect of selective attention. This was because previously presented but unattended strings were not used in the test phase in accordance with the convention of AG learning literature. This experiment was designed to clarify the ambiguity in the literature.

Experiment

Purpose

What is really learned in AG and how selective attention works in AG learning are worth investigating because they contributes not only to understanding the AG learning mechanism but also clarify the position of AG learning in implicit learning procedures represented by the SRT procedure.

Method

Participants Forty undergraduates from University of Tokyo participated in the experiment and received 500 yen following the completion of the experimental session. Assignments on attention conditions and GLOCAL strings were counterbalanced.

```
DDDDD    O  O      G   G
D        OO  O     G   G
D        O O O     G   G
D        O  OO     G   G
DDDDD    O  O        GGG
```

Figure 3: A sample of the GLOCAL strings used to convey instructions to the participants.

Stimuli The two AGs used were identical to those used by Tanaka et al. (2006) (Figure 1). These two AGs were less similar in their abstract structure according to the calculation based on the graph theory. In addition, they were constructed with the constraint that they did not share any chunks (e.g., VX), by using five letters: J, N, T, V, and X. Then, the letters B, F, L, Y, and Z were systematically substituted for J, N, T, V, and X in the second AG. This translation enabled each of these two AGs to hold unique letters as well as bigrams at the superficial level. In other words, the participants could not answer the question about one AG by using the learned knowledge of another. In fact, the results of the manipulation by Tanaka et al. revealed that the participants could not make correct judgments with regard to the grammar extracted from the unattended aspect of GLOCAL strings.

Eighteen grammatical strings were constructed using 3–6 letters from each AG to create the GLOCAL strings (e.g., Figure 2). Two types of series of GLOCAL strings were constructed from the strings following two AGs to counterbalance them between participants: one type of GLOCAL string followed Grammar 1 at the global aspect and Grammar 2 at the local aspect, whereas the other type followed Grammar 1 at the local aspect and Grammar 2 at the global aspect. GLOCAL strings were presented as white uppercase letters against a black background. Small letters were used in MS Gothic 12 pt. Seven small letters were arranged vertically to obtain one large letter. Eight small letters were arranged horizontally to obtain F, J, L, and X; nine, to obtain B, N, T, and Y; thirteen, to obtain V; and seven, to obtain Z. The height of a letter on the screen was about 3.2 cm and its width was about 1.8 to 3.0 cm. The distance between the display and participants was about 60 cm. These GLOCAL strings were used in the learning phase.

Twenty strings following each grammar were constructed for the test phase. These strings, composed of five or six letters, were not GLOCAL but regular letter strings; half of these were used in the learning phase and were referred to as presented grammatical strings. The remaining strings were not identical to any of the strings presented in the learning phase and were referred to as novel grammatical strings. All of these grammatical strings were used to construct nongrammatical strings that violated both grammars by placing one or two characters in nonpermissible locations. A constraint in the nongrammatical strings was that they were composed of the same letters as the grammatical ones.

Four types of string pairs were constructed for the test phase. The first type, "PG," paired a presented grammatical string at the global aspect of GLOCAL strings in the learning phase with a nongrammatical one based on the grammar extracted from the global aspect of the GLOCAL strings. The second type, "NG," paired a novel grammatical string at the global aspect of GLOCAL strings in the learning phase with a nongrammatical one based on the grammar that was extracted from the global aspect of the GLOCAL strings. Similarly, the third type was termed as "PL" and the fourth, as "NL." It should be noted that the difference between global and local was counterbalanced since half the participants were presented with GLOCAL strings that followed Grammar 1 at the global aspect and the other half, with the second type of GLOCAL strings that followed Grammar 1 at the local aspect in the learning phase. Each type comprised ten pairs. Matching pairs were randomized for each participant with the constraint that two strings should have the same length.

Design This was a $2 \times 2 \times 2$ mixed design. The first factor was Attention. Whether the participants were instructed to attend to the global or local aspect of the GLOCAL strings

in the learning phase was manipulated. This was a between-participants factor.

The second factor was Grammar. In the test phase, half of the pairs required the participants' judgment based on the grammar extracted from the global aspect of the GLOCAL strings, whereas the other half required their judgment based on the grammar extracted from the local aspect. This was a within-participant factor.

The last factor was Presentation, and it revealed whether the grammatical strings of the pairs used in the test phase were presented as the aspect of the GLOCAL strings in the learning phase. This was also a within-participant factor.
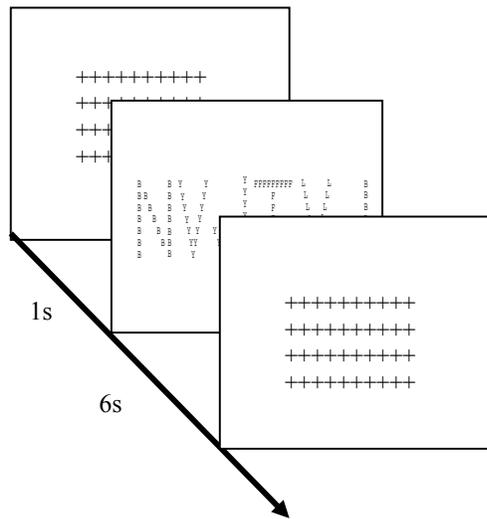

Figure 4: The flowchart of the learning phase. Note the white letters on the black background.

Procedure This experiment comprised learning and test phases. At the onset of the learning phase, the participants were informed that the stimuli comprising some large letters composed of small letters would be presented on the display. The participants were presented with a sample stimulus that was not used in our experiment (see Figure 3). This sample could be read as "dog" when the participants paid attention to the local aspect of the GLOCAL strings, and as "inu" (Japanese for "dog") when they attended to the global aspect.

The participants in the global attention condition were instructed to globally write the string on a sheet with ruled lines and read the sample as "inu"; the same instructions were provided for the presented series of GLOCAL strings (Figure 3). Conversely, the participants in the local attention condition were instructed to locally write the string on the sheet and read the sample as "dog"; the same instructions were provided for the presented strings.

During the learning phase, 18 GLOCAL strings were presented on the display for 6 s, with each GLOCAL string

being presented six times. A mask stimulus comprising many "+" signs in the area where the GLOCAL strings were intended to be displayed was presented for 1 s during the interval between the presentation of two GLOCAL strings. The learning phase was followed by the test phase.

In the test phase, the participants were informed that two strings would be presented in the upper and lower regions of the display; that the GLOCAL strings followed some rules; and that while one string of a pair followed these rules, the other did not. They were required to press either of two specific keys to identify the strings. Forty pairs were presented to each participant in random order. A pair of strings remained on display until each participant pressed one key. The presentation of the strings of a pair in the upper region was also randomized for each participant with the constraint that the grammatical strings of pairs be presented an equal number of times in each region. The 40 pairs were presented twice; in other words, the participants had to provide answers for 80 pairs.
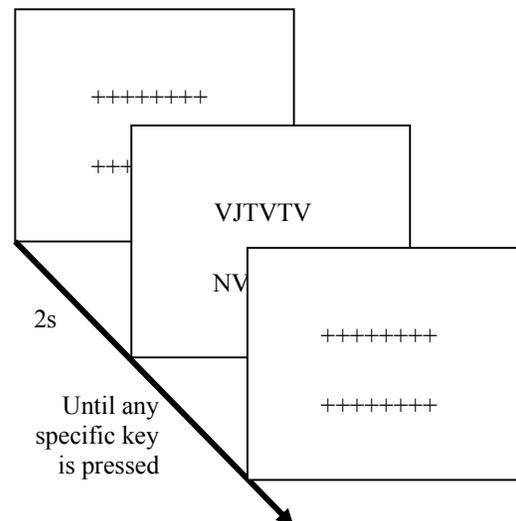

Figure 5: The flowchart of the test phase. Note the white letters on the black background.

Results

Figure 6 shows the means of the selection rates of the grammatical strings of each condition in the test phase. All chance levels were 0.5: if the participants could not acquire any knowledge of each grammar, the selection rates of the grammatical strings were 0.5 in the test phase.

First, the selection rates of the grammatical strings were submitted to a 2 × 2 × 2 mixed analysis of variance (ANOVA) with Attention (attended to the global or local aspects of the GLOCAL strings in the learning phase: between participants) and Grammar (extracted from the global or local aspect of the GLOCAL strings) and Presentation (presented or novel as an aspect of the

GLOCAL strings in the learning phase). The main effect of Attention was significant, $F(1,38) = 6.06$, MSE $= 0.0295$, $p < .05$. The mean selection rate of grammatical strings at the global attention condition (0.63) was higher than that at the local attention condition (0.56). The main effect of Grammar was also significant, $F(1,38) = 9.31$, MSE $= 0.0185$, $p < .01$.
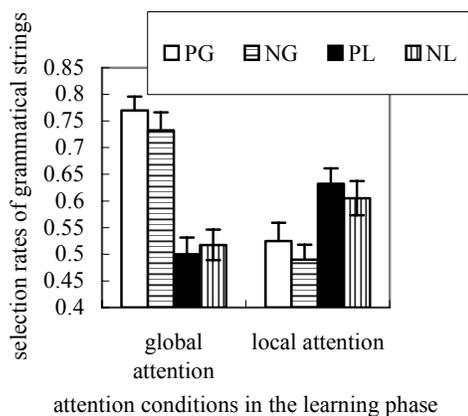

Figure 6: Means (±SE) of the selection rates of grammatical strings of four types of pairs: presented global (PG), novel global, (NG), presented local (PL), and novel local (NL).

The selection involved in the global aspect of GLOCAL strings (the mean selection rate was 0.63) was more correct than that in the local aspect (0.56). The interaction between Attention and Grammar was significant, $F(1,38) = 67.62$, MSE $= 0.0185$, $p < .01$. The results of the simple main effect revealed that the effects of Attention influenced the global and local grammars, $F(1,76) = 49.50$ and 10.08, respectively, MSE $= 0.0240$, $p < .01$. Similarly, the results of the simple main effect revealed that the effect of Grammar influenced the attended aspect as compared to the global and local aspects, $F(1,38) = 63.55$ and 13.38, respectively, MSE $= 0.0240$, $p < .01$. In the global attention condition, the mean correct rate of global grammar was significantly higher than of local grammar, whereas the relationship between global and local grammars with regard to the mean correct rates was inversed in the local attention condition. The main effect and interaction with respect to Presentation factor were not significant ($Fs < 1$).

In order to examine whether the participants can learn AG from the unattended aspect of GLOCAL strings, using t tests (two-tailed), we compared the selection rates of the grammatical strings with the chance level in each condition. However, there was no significant difference with regard to the chance level in the four conditions ($ts < 1$).

Discussion

The interaction between Attention and Grammar reveals that the participants could learn AG more from the attended aspect than from the unattended one. In fact, there are no differences between unattended selection rates and the chance level in all conditions. This result shows that selective attention is necessary for AG learning.

What Is Really Learned in AG Learning

In terms of the comparison between presented and novel strings, there is no significant difference in the main effect and interactions with regard to Presentation. This suggests that in the test phase, the participants judged grammaticality based on their grammatical knowledge rather than their representation of presented strings. This may be interpreted as counterevidence against the viewpoint that participants could not extract AG as the abstract structure from the presented strings (e.g., Kinder, Shanks, Cock, & Tunney, 2003). This result cannot deny the existence of the representation of presented strings in memory. It is possible to consider that participants' representation of presented strings will be generalized through repeated presentation. Conceivably, the extraction of AG may precede the storage of presented strings. In any case, as per the objective interpretation of this result, participants' behavior could not be interpreted solely based on the trace of presented strings since there was no significant difference between the presented and novel grammatical strings.

In the context of the mere exposure effect (using preference judgment instead of grammaticality judgment), some literatures discuss the difference between presented and novel grammatical strings. For instance, in the condition of relatively short time presentation (Newell & Bright, 2003) or in one using low familiarity characters (Zizak & Reber, 2004), the preferences for novel grammatical strings disappeared, whereas those for presented grammatical strings still remained.

However, in this experiment, there was no significant difference between presented and novel grammatical strings, whereas there was a clear difference between grammatical and nongrammatical strings. This is because in this experiment, grammatical judgment was used as a more direct measurement to investigate the existence of AG than was preference judgment. On the whole, as far as observed behavior is concerned, the finding that participants can learn AG from presented strings seems to be appropriate for the interpretation of their behavior in the AG learning procedure. Future research on this topic should control for similarities between novel grammatical strings and nongrammatical strings based on some outer standard, although it is difficult to define the similarity between strings with psychological validity.

Relationship between Implicit Learning and Selective Attention

With regard to selective attention, the present results reveal that the participants could not select the presented but unattended strings as well as the novel and unattended ones. This is interesting because it is known that during the preattention process, the unattended aspect of compound letters interferes with the processing of the attended aspect. This finding can be interpreted as evidence for input processing independent of the subjective intention for seeing. Some studies in the context of the mere exposure effect have also revealed the existence of representation in memory of presented strings themselves. Given such findings, the present results could postulate that information on the unattended aspect of compound letters is inevitably processed and stored as its representation in memory. Thus, the present results can be interpreted as evidence that for experimental tasks, the participants could inhibit the information of the unattended aspect. Note that participants were not instructed that the presented stimuli shared some rules in the learning phase. Therefore, these processes were not modulated by participants' conscious intentions. It could be interpreted that for the sake of selective attention, a series of attended strings were generalized and other series of unattended strings were inhibited.

The finding that the acquisition of very complex rules is caused by only controlling what to see and not orienting the goal as the purpose of behavior is suggestive. It is possible to consider that the learning systems investigated using the implicit learning paradigm are the same as the substrate of the acquisition of various skills, including the acquisition of a mother tongue without a definite goal.

Furthermore, with regard to the selection rates of the grammatical strings for the attended grammar, the rates in the global attention condition were higher than those in the local attention condition. This global advantage propensity is important since it is known that global processing antecedes local processing in various contexts (Navon, 2003, for a review). The asymmetry based on different aspects of the GLOCAL strings may suggest that AG learning, which seems higher cognitive processing than visual processing, also depends on the input process of visual information processing. In fact, participants in both the conditions could write almost all strings of the attended aspect of the GLOCAL strings regardless of global or local conditions. The perceptual load theory (Lavie, 1995) enables us to explain this asymmetry. The difference in the correct rates of the grammatical judgment task between the global and local attention conditions can be interpreted as the difference of perceptual load between two visual aspects. The ease of perception of the global aspects compared with that of the local aspects could cause the higher correct rates of the global condition than the local condition at the attended aspects. It seems that the interaction between selective attention and visual input as the environment variable also plays an important role in AG learning.

References

Buchner, A., & Wippich, W. (1998). Differences and commonalities between implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning.* Thousand Oaks, CA: Sage.

Hibi, Y., Takeda, Y., & Yagi, A. (2002). Attention level and negative priming in hierarchical patterns. *Japanese Psychological Research*, *44*, 241-246.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *The Psychonomic Bulletin & Review*, *12*, 100-106.

Kinder, A., Shanks, D. R., Cock, J., & Tunney, R. J. (2003). Recollection, fluency, and the explicit/implicit distinction in artificial grammar learning. *Journal of Experimental Psychology: General*, *132*, 551-565.

Kramer, R. S. S., & Parkinson, B. (2005). Generalization of mere exposure to faces viewed from different horizontal angles. *Social Cognition*, *23*, 125-136.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human, Perception, and Performance*, *21*, 451-468.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353-383.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica*, *114*, 273-309.

Newell, B. R., & Bright, J. E. H. (2003). The subliminal mere exposure effect does not generalize to structural related stimuli. *Canadian Journal of Experimental Psychology*, *57*, 61-68.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, *115*, 163-196.

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

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *77*, 317-327.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219-235.

Tanaka, D., Kiyokawa, S., Yamada, A, & Shigemasu, K. (2006). Role of Selective Attention in Artificial Grammar Learning. *The 28th Annual Conference of Cognitive Science Society*, 2615.

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.

Zizak, D. M., & Reber, A. S. (2004). Implicit preferences: The role(s) of familiarity in the structural mere exposure effect. *Consciousness and Cognition*, *13*, 336-362.