



Universiteit
Leiden
The Netherlands

Implicit artificial grammar learning: effects of complexity and usefulness of the structure

Bos, E.J. van den

Citation

Bos, E. J. van den. (2007, June 6). *Implicit artificial grammar learning: effects of complexity and usefulness of the structure*. Department of Cognitive Psychology, Leiden University Institute for Psychological Research, Faculty of Social Sciences, Leiden University. Retrieved from <https://hdl.handle.net/1887/12037>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12037>

Note: To cite this publication please use the final published version (if applicable).

Chapter 2

Effects of grammar complexity

Abstract

Implicit learning has been viewed as a process that automatically abstracts any regularity present in the environment. The process would be especially suited for complex regularities that are difficult to figure out intentionally (Reber, 1989). The present study investigated whether or not implicit learning of artificial grammars is influenced by the complexity of the grammars. In addition, it was examined whether complexity affected the type of knowledge acquired. Ten artificial grammars of varying complexity were studied either implicitly or explicitly. The results showed better performance after implicit than after explicit learning. However, implicit learning was hampered by increasing complexity. In addition, there was an effect of complexity on the type of knowledge acquired: first-order dependencies were learned for all grammars, whereas the acquisition of second-order dependencies was negatively affected by complexity. These results contrast with an automatic abstraction view of implicit learning.

Introduction

When people are presented with exemplars of a simple structure, they may intentionally try to grasp the regularities and become aware of the knowledge they acquire. In addition to this explicit way of learning, Reber (1976, 1989) proposed that structures can also be learned implicitly; without any intention to learn and without complete awareness of the acquired knowledge. To observe implicit structure learning, a minimum level of complexity of the stimuli is required (Lee, 1995; Reber, 1976). Therefore, complex artificial grammars have been introduced as a means to study implicit learning (Reber, 1976, 1989).

Reber (1976) also claimed that implicit learning is especially suited to complex structures. While explicit learning would be hampered by increasing complexity, the implicit learning process would continue to automatically abstract any structure present in the environment. This counterintuitive suggestion has gone virtually unaddressed in the literature on artificial grammar learning (AGL). The

present study will investigate whether or not AGL is affected by the complexity of the grammar. In addition, the influence of complexity of the grammar on the type of knowledge acquired will be examined.

Learning and complexity

Various studies have indicated that implicit and explicit learning are differentially affected by the complexity of the structure to be learned. For example, Reber (1976) found that a complex artificial grammar could better be learned implicitly than explicitly. In the induction phase of the experiment, the implicit learning group was instructed to memorize letter strings without being informed that they were generated by a grammar. The explicit learning group received the additional instruction to look for rules underlying the exemplars. Subsequently, when all participants had been informed about the existence of the grammar, they were instructed to judge whether or not new exemplars were grammatical. Participants who had learned implicitly were more often correct than participants who had learned explicitly.

For simple materials, in contrast, there are some indications that implicit learning is less efficient than explicit learning. Mathews et al. (1989) showed that a finite state grammar could best be learned implicitly, whereas explicit learning was more successful with a biconditional grammar. Johnstone and Shanks (2001) replicated this finding and suggested that the different results for the two types of grammar might be due to differences in complexity. They argued that biconditional grammars are less complex than finite state grammars, since they consist of a smaller number of rules. In addition, Reber (1989) pointed out that, in the few studies in which explicit learning of a finite state grammar led to better results than implicit learning, the complex pattern had been structured for the participants. Regularities were made salient by making the pattern meaningful or by presenting together those strings that followed the same rule. This suggests that explicit learning is successful when complexity is reduced.

These AGL-studies suggest that implicit learning is better suited to complex structures than explicit learning, whereas explicit learning is more efficient for simple materials. In addition, Reber (1993) noted in a review of his work that variations in the complexity of finite state grammars above the level required to observe implicit learning have little effect on performance. However, these studies do not support firm conclusions on the effects of structure complexity on implicit and explicit learning,

because complexity has not been systematically manipulated. The lack of such a study in the domain of artificial grammar learning is somewhat surprising, as the possibility that highly complex patterns can be learned implicitly has been questioned by researchers on natural grammar learning (Gold, 1967; Pinker, 1989).

In the Serial Reaction Time (SRT) paradigm (Nissen & Bullemer, 1987), several studies have manipulated the complexity of the structure participants were exposed to. Some of these studies have indeed shown that explicit learning is hampered by increasing complexity of the structure, while implicit learning is not. When a fixed sequence determined the successive locations on a computer screen where a stimulus appeared, the reaction times of participants who had to press a key corresponding to the stimulus' location decreased irrespective of whether or not they performed the additional task of trying to discover the sequence. However, when complexity of the sequence was increased by alternating predictable locations with random ones, only participants who were not trying to discover the sequence demonstrated learning (Fletcher et al., 2005). In an experiment of Reed (reported in Reed & Johnson, 1998) the location of each stimulus depended on the background colors of the locations. This relationship was determined by one, two or four rules. Explicit learning decreased with increasing complexity of the relationship, whereas implicit learning did not.

However, other SRT-studies provided evidence that implicit learning, too, is negatively affected by increasing complexity. For example, Stadler (1992) used sequences with various levels of statistical structure and found that the implicit learning effect decreased with decreasing predictability of the sequences. Similarly, a study by Soetens, Melis and Notebaert (2004) showed more implicit learning for sequences containing first-order dependencies than for sequences containing at least second-order dependencies. The contrasting findings of these SRT-studies may well be due to differences in the definitions and manipulations of complexity that have been used. As it is, these studies do not allow an unambiguous conclusion on the relationship between complexity and implicit learning. Therefore, the present study will further investigate the effects of the complexity of a structure on implicit and explicit artificial grammar learning.

Measures of complexity

As noted above, previous studies on implicit learning have used various ways to measure the complexity of the structure that had to be learned. Firstly, a relatively

simple measure of the complexity of artificial grammars can be found in the work of Johnstone and Shanks (2001). They used the number of rules a grammar consists of as an indication of its complexity. This number is equal to the number of letters that could be added to a string at each transition between internal nodes of the grammar plus the number of letters that can terminate a string at each end node (Johnstone & Shanks, 1999). If the end node of the grammar is considered to be equivalent with the initial node, this comes down to the number of links in the grammar. As discussed before, the number of rules has also been used by Reed (reported in Reed & Johnson, 1998) to vary the complexity of the regularities that had to be learned in an SRT-task.

At another level of description, Perruchet and Vinter (1998) noted that learning may be influenced by the number of associations that has to be acquired. For example, a structure that can be described using a few bigrams (i.e. two-letter chunks) repeatedly is simpler than a structure that can only be described by many unique bigrams. The number of bigrams required to describe a structure can therefore be viewed as a second measure of complexity.

A third measure can be found in the SRT-paradigm, where the complexity of a structure is often related to its predictability (e.g. Cohen, Ivry, & Keele, 1990; Reed & Johnson, 1994; Soetens et al., 2004; Stadler, 1992). In the simplest case, a stimulus at a certain location on the screen is always followed by a stimulus at one other location. Knowing the location of the present stimulus is then sufficient to predict where the next will appear. The structure becomes more complex when the next stimulus can appear at several locations, depending on the location of the stimulus before the present one. In this way, the number of previous elements required to predict the next can be used to measure the complexity of a structure.

Finally, Boltt and Jones (2000) proposed topological entropy as a specific measure of the complexity of finite state grammars. A detailed explanation of the computation of topological entropy is provided in Appendix A. In short, the complexity of an artificial grammar is defined as the growth rate of the number of unique exemplars of a given length that the grammar generates as string length goes to infinity (Boltt & Jones, 2000). The authors explain that, for most grammars, the number of different exemplars that can be generated grows exponentially as exemplars of greater length are considered. Topological entropy is a measure of the exponent with which the number of possible exemplars increases. Put simply: the

complexity of a grammar increases according to this measure as it generates more different exemplars of any given length.

Although the measures discussed above are sensitive to different aspects of complexity, they are highly inter-correlated. This is because each measure is sensitive to different effects of adding links to a grammar. Adding a link by definition increases the number of rules and it can also allow for the generation of additional unique exemplars, increasing topological entropy. At the same time, the new link may create a new bigram, increasing the number of bigrams required to describe the structure. Alternatively, the new link may cause an existing bigram to occur at another location in the grammar. This may increase the number of previous elements that have to be taken into account to predict which letters can succeed the bigram. (See Appendix A for an additional explanation of the correlation between topological entropy and the number of elements required to predict the next). Table 1 illustrates the inter-correlatedness of the measures. The correlations are based on the scores of the ten artificial grammars used in this study on each of the four measures of complexity (see Table 2 on page 16).

Table 1

Correlations and partial correlations between four measures of complexity

	NR		NB		NE		TE	
	<i>r</i>	partial <i>r</i>	<i>r</i>	partial <i>r</i>	<i>r</i>	partial <i>r</i>	<i>r</i>	partial <i>r</i>
NR	1.000	1.000	.994**	.926**	.520	-.471	.718*	.418
NB			1.000	1.000	.570	.126	.757*	-.058
NE					1.000	1.000	.965**	.993**
TE							1.000	1.000

Note. NR = number of rules; NB = number of bigrams; NE = number of elements required to predict the next; TE = topological entropy.

* $p < .05$. ** $p < .01$.

The partial correlations in Table 1, indicating the correlation between two measures when the contributions of the other two measures have been removed, reveal that the four measures of complexity, in fact, measure two different aspects. On the one hand, the number of rules and the number of bigrams seem to measure the number of associations that have to be learned. The number of elements required to

predict the next and topological entropy, on the other hand, seem to measure the length of the dependencies that have to be learned. The number of bigrams and topological entropy are more fine-grained than their counterparts. As they are probably the most sensitive, these two measures will be used in the present study to investigate whether or not these aspects of complexity affect artificial grammar learning.

The type of knowledge acquired in AGL

Interestingly, one of the aspects of complexity that will be examined in this study, the length of the dependencies that have to be learned, is often discussed in the context of the type of knowledge that is actually acquired in AGL. There is an ongoing debate on whether participants learn first- (bigrams), second- (trigrams) or higher-order dependencies. In this debate, there is general agreement that participants can learn first-order dependencies. Kinder (2000) showed that participants' grammaticality judgments of different types of test exemplars mainly depended on the legality of the bigrams in the exemplars. In addition, Perruchet and Pacteau (1990) showed that knowledge of bigrams could be sufficient to explain performance in an AGL-experiment. Participants who studied a list of bigrams in the induction phase performed as well on the grammaticality judgment task as participants who studied complete exemplars.

However, Gomez and Schvaneveldt (1994) showed that participants who studied complete exemplars acquired more knowledge than participants who studied bigrams. In their experiment, half of the ungrammatical test exemplars contained an illegal bigram and the other half contained a legal bigram at an illegal location (i.e. an illegal trigram). Participants who had studied bigrams only recognized exemplars with an illegal bigram as ungrammatical, whereas participants who had studied complete exemplars also discarded illegal trigrams, demonstrating knowledge of second-order dependencies.

Acquisition of even higher-order dependencies has been shown in studies controlling for chunk strength: i.e. the frequency of occurrence in the induction phase of the bigrams and trigrams composing the test exemplars (Knowlton & Squire, 1994). In studies lacking this control, grammaticality judgments can both be based on the familiarity of bigrams and trigrams and on knowledge of higher-order dependencies. Studies in which chunk strength and grammaticality were varied orthogonally or in which chunk strength was equated for grammatical and

ungrammatical strings have shown that both types of knowledge are acquired (Knowlton & Squire, 1996; Lieberman, Chang, Chiao, Bookheimer & Knowlton, 2004; Meulemans & Van der Linden, 2003).

The research discussed above indicates that artificial grammar learning can result in knowledge of first-, second- and higher-order dependencies. However, some studies suggest that longer dependencies are more difficult to learn implicitly. For example, Meulemans and Van der Linden (1997) showed that participants acquired knowledge of higher-order dependencies when they were presented with a large number of exemplars, whereas they only acquired knowledge of bigrams and trigrams from presentation with a small number. In addition, there is some evidence that only first-order dependencies can be learned without awareness. Participants who were sensitive to violations of second-order dependencies on a grammaticality judgment test also showed explicit knowledge of trigrams on a recognition test (Gomez, 1997). If implicit learning is hampered by increasing complexity and higher-order dependencies are indeed more difficult to learn, then we would expect this type of knowledge to be particularly affected by increasing complexity.

In summary, the present study will address two questions. Firstly, we will investigate the relationship between complexity of the grammar and performance on an AGL-task under implicit and explicit learning instructions. Secondly, we will examine whether complexity of the grammar influences the acquisition of first- and second-order dependencies. We predict that, if implicit learning is affected by complexity, the negative effect will be strongest for second-order dependencies.

Method

Participants

Sixty-one undergraduate students of Leiden University (20 male, 41 female, 18-41 years of age) participated in the experiment. They received either course credits or € 4.50 for their participation. The data from one participant had to be discarded, because, as a non-native speaker of Dutch, he turned out to be unable to understand the instructions.

Design

The experiment consisted of an induction phase and a test phase. In both phases, exemplars were presented in random order. There were three independent variables. Firstly, the instruction for the induction phase was varied between

participants. One half of the participants was instructed to memorize the exemplars. The other half was informed that the exemplars had been generated according to certain rules and received the instruction to search for these rules.

A second between-subjects variable was the complexity of the grammar, as measured by its number of bigrams and its topological entropy. Ten artificial grammars were used to vary complexity. This large number was chosen, because the research questions were directed at a continuous relationship between grammar complexity and performance, rather than at a difference in performance on two particular grammars. The ten artificial grammars were each studied by three participants under each instruction, providing 30 data points per group for the regression analyses.

Finally, type of test exemplar was varied as a within-subjects factor to examine the type of knowledge participants acquired. Apart from grammatical exemplars, there were different types of ungrammatical exemplars containing various violations. The relative importance of knowledge of first- and second-order dependencies in detecting the violations varied for the different types of exemplars. Therefore, differences in performance on these types of exemplars indicated differences in the extent to which knowledge of first and second-order dependencies had been acquired. The dependent variable was the mean proportion of test exemplars correctly classified as grammatical or ungrammatical. In addition, participants provided confidence ratings for their grammaticality judgments.

Table 2

Scores on four measures of complexity for each of the grammars used in this study

Grammar	NR	NB	NE	TE
a	20	30	2	.5543
b	21	33	3	1.2038
c	22	36	3	1.3506
d	22	36	4	2.0496
e	23	38	2	.7131
f	24	41	4	2.2454
g	24	40	3	1.5173
h	25	44	4	2.4122
i	26	45	3	1.6898
j	27	47	4	2.5761

Note. NR = number of rules; NB = number of bigrams; NE = number of elements required to predict the next; TE = topological entropy.

Materials

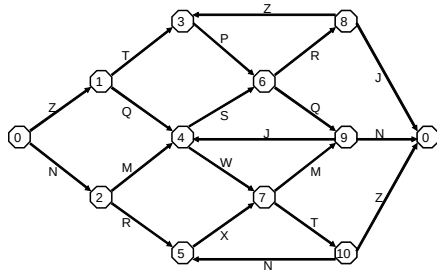
The stimuli in this experiment were exemplars from ten finite state grammars (See Figure 1 on page 18). All grammars used the same letters: J, M, N, P, Q, R, S, T, W, X, Z and consisted of the same 11 states. However, these states were connected in different ways and by a varying number of links to produce different scores on the complexity measures. The scores on four complexity measures are provided in Table 2 for each of the grammars.

A computer program generated a set of 120 unique exemplars for each grammar. Each exemplar was generated by moving along the arrows from the initial state (0) to the end state (0) of the graph, while adding the corresponding letter to the exemplar. The length of the exemplars varied between 5 and 11 letters. Of each set, 60 exemplars were assigned to the induction phase and 50 to the test phase, balanced over the paths of the grammar. Five of the remaining exemplars were used on practice trials in the test phase; the rest was discarded. The stimuli for practice trials in the induction phase consisted of number strings unrelated to the grammars.

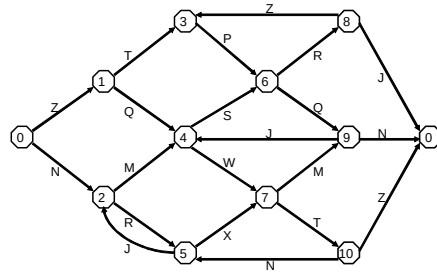
The 50 exemplars assigned to the test phase were subdivided into two subsets. One subset consisted of unaltered grammatical exemplars, whereas the other consisted of exemplars that were made ungrammatical by switching two adjacent letters; excluding the first and last. Switching letters could result in a violation of a first-order dependency: a bigram that cannot occur according to the grammar. Alternatively, it could result in a violation of a second-order dependency: a trigram that cannot occur according to the grammar. Switching two inner letters of an exemplar always affects three transitions, thus producing one of four possible combinations of violations: three first-order violations, two first and one second-order violation, one first and two second-order violations or three second-order violations. These combinations were introduced into the ungrammatical test sets according to their probability of occurrence in the complete set of 120 strings generated by each grammar. As the probability of three second-order violations was very low in most grammars, only three types of ungrammatical exemplars were discerned in this study: exemplars with 3 first-order violations, exemplars with 2 first and 1 second-order violation and exemplars with 2 or more second-order violations.

All stimuli were displayed on a computer monitor as black text (Arial 18, bold) against a white background. Participants were seated in front of the computer monitor at a distance of about 50 cm. They reacted by pressing keys on a keyboard.

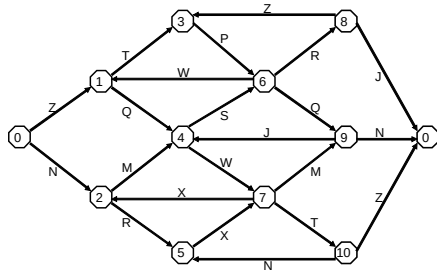
Grammar a



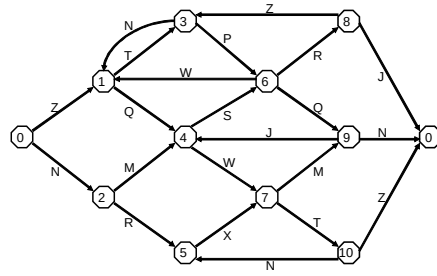
Grammar b



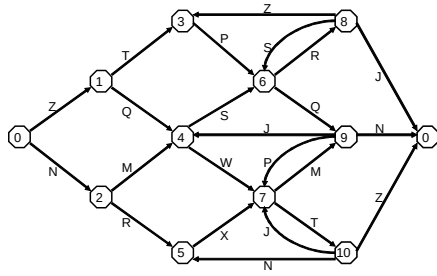
Grammar c



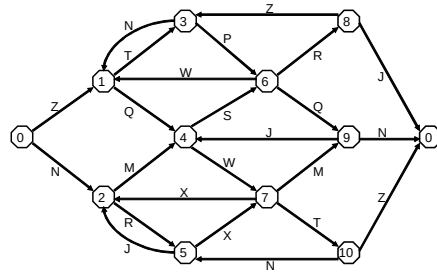
Grammar d



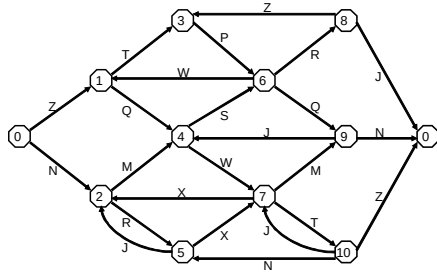
Grammar e



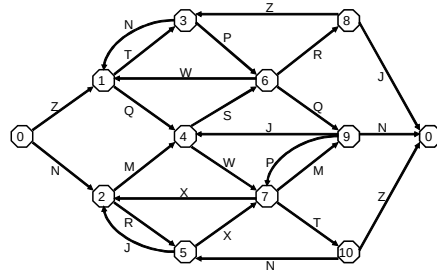
Grammar f



Grammar g



Grammar h



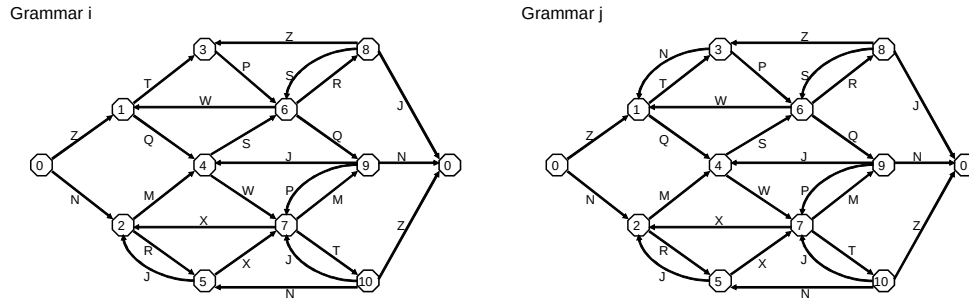


Figure 1. The ten artificial grammars used to generate the stimuli for this study.

Procedure

Participants were tested individually in a dimly lit test booth. At the beginning of the experiment, they were informed that it would consist of two parts and would involve studying letter strings for a subsequent test. One half of the participants received the additional information that the exemplars were generated according to a complex set of rules. Their instruction was to search for these rules. The other participants were only told to memorize the individual exemplars.

All participants were presented with 5 practice trials and 60 experimental trials. Each trial started with a fixation cross appearing in the middle of the screen. After 1 second the cross was replaced by an exemplar centered at the fixation point. The exemplar was displayed for 5 seconds. Then participants with the ‘memorize’-instruction were prompted to reproduce the exemplar. Participants with the ‘look for rules’-instruction were prompted to enter a part of the exemplar that they thought was important. After pressing the enter-key, participants were again presented with the original exemplar for 2 seconds so that they could check their answer. Finally, the screen turned blank for 1 second before the next trial began.

In the second part of the experiment, all participants were informed that the letter strings presented in the first part had been generated according to a complex set of rules. They were instructed to judge whether or not new letter strings followed the same rules. Participants were required to press the ‘j’-key (for ‘ja’: ‘yes’) if they thought that an exemplar followed the rules and the ‘n’-key (for ‘nee’: ‘no’) if they thought that an exemplar did not follow the rules. In addition, they were required to indicate their confidence in each judgment on a scale from 1 (very little) to 5 (very much) by pressing one of the number keys on the keyboard.

Participants were presented with 5 practice trials followed by 50 experimental trials. Each trial began with a fixation cross appearing in the middle of the screen. After 1 second the cross was replaced by an exemplar centered at the fixation point. When the participant pressed the 'j' or 'n'-key, the screen turned blank for 1 second. Subsequently, the confidence scale was presented until the participant pressed a number from 1 to 5. A final blank screen separated two consecutive trials by 1 second. When participants had completed all trials, they were thanked for their participation. The experiment took about 30 minutes.

Analyses

The effect of complexity on performance in AGL. Linear regression analyses (enter method) were performed to determine whether a grammar's complexity affects how well artificial grammars are learned. The dependent variable was the proportion of correct classifications in the test phase; the predictors were the two measures of complexity: number of bigrams and topological entropy. The three regression models that can be built with two predictors (including number of bigrams, including topological entropy, and including both) were compared to determine how the proportion of correct classifications could best be predicted. Separate analyses were performed for the group instructed to memorize the exemplars and the group instructed to look for the underlying rules. An independent samples *t*-test was performed to examine the effect of instruction in the induction phase on the proportion of correct classifications in the test phase over all grammars.

To examine whether or not the instruction to memorize exemplars had induced implicit learning, the 'guessing criterion' (Dienes & Altmann, Kwan & Goode, 1995) was used. Dienes and Berry (1997) proposed that knowledge can be implicit in the sense that people are unaware of possessing it, i.e. when meta-knowledge is lacking. According to the guessing criterion, people possess implicit knowledge if they perform significantly above chance while they claim to be guessing. Knowledge classified as explicit according to this criterion has been shown to be affected by divided attention, while knowledge classified as implicit was not (Dienes et al., 1995), suggesting that the criterion makes a meaningful distinction between implicit and explicit knowledge. Therefore, the proportion of correct grammaticality judgments was computed for trials on which participants had rated (very) little confidence in their judgment. Participants' scores were submitted to a one-sample *t*-test. Scores that were based on less than three trials were excluded from the analysis.

The effect of complexity on the type of knowledge acquired in AGL. To determine whether or not complexity affects the type of knowledge acquired in artificial grammar learning, separate linear regression analyses were performed for each type of test exemplar (grammatical, 3 first-order violations, 2 first and 1 second-order violation, 2 or more second-order violations). The dependent variable was the proportion of correct classifications in the test phase; number of bigrams and topological entropy were the predictors. Again, participants' scores were only included in the analysis if they were based on more than two trials. The analyses were performed separately for the group instructed to memorize the exemplars and the group instructed to look for the underlying rules.

Results

The effect of complexity on performance in AGL

For the group instructed to look for rules underlying the exemplars, none of the regression models was significant ($p \geq .272$), indicating that neither the number of bigrams nor topological entropy was a good predictor of the overall proportion of correct classifications in the test phase. For the group of participants who had been instructed to memorize exemplars in the induction phase, in contrast, all models reached significance. The model including only the number of bigrams accounted for 14.1% of the variance in the overall proportion of correct classifications ($F(1,28) = 4.601, p = .041$). The model only including topological entropy accounted for 19.8% of the variance ($F(1,28) = 6.931, p = .014$) and the model including both predictors accounted for 20.2% of the variance ($F(2,27) = 3.415, p = .048$). Adding the number of bigrams was no significant improvement to the model including only topological entropy, however ($F_{change}(1,27) = .118, p = .734$), indicating that topological entropy was, by itself, the best predictor of proportion correct in the test phase. The negative relationship is illustrated by Figure 1 on page 22, which shows the predicted and observed scores per instruction for each measure of complexity.

One sample t -tests on the participants' scores over all trials indicated that the mean proportion correct (see Table 3 on page 23) was significantly above chance, both for participants who had memorized exemplars ($t(29) = 8.764, p < .001$) and for participants who had looked for the underlying rules ($t(29) = 4.305, p < .001$). However, the proportion correct was higher after memorizing than after looking for rules ($t(58) = 3.126, p = .003$).

For trials on which participants had rated little confidence in their judgments, one-sample t -tests showed that the proportion correct was significantly above chance for participants who had memorized exemplars ($M = .627$, $SD = .169$; $t(25) = 3.847$, $p = .001$), but failed to reach significance for participants who had looked for rules ($M = .554$, $SD = .158$; $t(26) = 1.775$, $p = .088$). So, participants in the memorize group acquired a significant amount of implicit knowledge, which could not be demonstrated in the look for rules group. This suggests that the two instructions succeeded in inducing different learning modes.

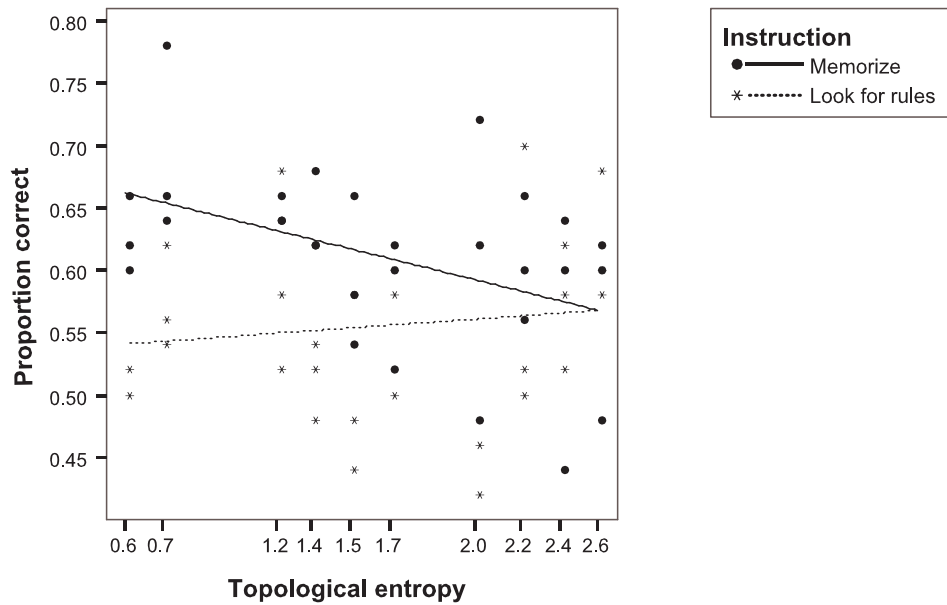


Figure 1. Predicted and observed proportions of correct grammaticality judgments after memorizing and looking for rules in the induction phase as a function of topological entropy.

The effect of complexity on the type of knowledge acquired in AGL

For the memorize group, none of the regression models was significant for the grammatical exemplars ($p \geq .407$), the exemplars with 3 first-order violations ($p \geq .395$) or the exemplars with 2 first and 1 second-order violation ($p \geq .115$). For the proportion of exemplars with 2 or more second-order violations, however, the regression model including only topological entropy reached significance ($F(1,22) = 5.885$, $p = .024$), accounting for 21.1% of the variance in the proportion of correct

classifications. The relationship was negative, which suggests that participants in the memorize condition recognized exemplars containing several illegal trigrams as ungrammatical when they worked with simple grammars, but lost this ability as complexity increased. For the other types of test exemplars, one sample *t*-tests indicated that overall performance was significantly above chance on grammatical exemplars ($t(29) = 9.280, p < .001$) and exemplars with 3 first-order violations ($t(29) = 5.911, p < .001$), but not on exemplars with 2 first and 1 second-order violation ($t(29) = -.396, p = .695$). The mean proportions of correct grammaticality judgments for each type of exemplar are shown in Table 3.

Table 3

Mean proportions correct and standard deviations for different types of exemplars by instruction.

	Memorize		Look for rules	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All strings	.612**	.070	.555**	.070
Grammatical	.672**	.102	.587*	.148
3 first-order violations	.673**	.161	.612*	.187
2 first, 1 second-order violations	.488	.170	.507	.146
2 or more second-order violations	.520	.220	.411	.257

Note. * $p < .05$. ** $p < .01$.

For participants who had looked for rules underlying the exemplars in the induction phase, none of the regression models was significant for the grammatical exemplars ($p \geq .797$), the exemplars with 2 first and 1 second-order violation ($p \geq .096$) or the exemplars with 2 or more second-order violations ($p \geq .732$). One sample *t*-tests indicated that overall performance (see Table 3) was above chance on grammatical exemplars ($t(29) = 3.201, p = .003$), but not on exemplars with 2 first and 1 second-order violation ($t(29) = .244, p = .809$) and exemplars with 2 or more second-order violations ($t(23) = -1.693, p = .104$). For exemplars containing 3 first-order violations, however, all regression models reached significance. The model including only the number of bigrams accounted for 22.6% of the variance ($F(1,28) = 8.196, p = .008$), the model only including topological entropy accounted for 25.6% of the variance ($F(1,28) = 9.620, p = .004$) and the model including both predictors

accounted for 27.6% of the variance ($F(2,27) = 5.148, p = .013$). However, adding the number of bigrams was no significant improvement to the model including only topological entropy ($F_{change}(1,27) = .759, p = .391$). Interestingly, the relationship between topological entropy and proportion correct was positive, indicating that more exemplars with 3 first-order violations were recognized as ungrammatical by participants who had looked for rules as complexity increased.

Discussion

The effect of complexity on performance in AGL

The first aim of this study was to examine whether or not performance on an AGL-task would be affected by increasing complexity of the grammars. Two different learning modes were induced by instructing participants either to memorize exemplars (implicit grammar learning) or to look for rules underlying exemplars (explicit grammar learning). The effect of increasing complexity was examined for both learning modes to test Reber's (1993) suggestion that explicit learning would be hampered by increasing complexity, while implicit learning would not.

After explicitly looking for rules in the induction phase, participants were overall less accurate in judging the grammaticality of new exemplars than after memorizing. In addition, explicit learning was not affected by complexity of the grammar, irrespective of whether it was measured by its number of bigrams or by its topological entropy. As previous studies indicated that explicit learning was more efficient than implicit learning for simple materials (Johnstone & Shanks, 2001; Lee, 1995), this suggests that an improvement in performance may occur at lower levels of complexity. For the range examined in the present study, performance after explicit learning was significantly above chance, but lower than after implicit learning. The latter finding is in line with Reber's (1976) claim that complex materials can better be learned implicitly than explicitly. However, support for this claim is refined by the findings discussed below.

For implicit learning, the results of the present experiment showed that performance on the test task was affected by complexity of the grammar, as measured by its number of bigrams and topological entropy. The regression analysis indicated a significant negative relationship: an increase in complexity caused the proportion of correct grammaticality judgments to decline. On the most complex grammar, implicit learning did not lead to more correct grammaticality judgments than explicit learning.

Although the findings of the present AGL-study contrast with some results obtained with the SRT-paradigm (Fletcher et al., 2005; Reed & Johnson, 1998), they are in line with the results of other SRT- studies (Stadler, 1992; Soetens et al., 2004), providing converging evidence that implicit learning is affected by complexity. These findings are at odds with a view of implicit learning as a process that automatically abstracts any structure present in the environment. Rather, implicit learning, like explicit learning, seems to be limited in scope. While structures of moderate complexity may well be acquired implicitly, the process is increasingly hampered in acquiring structures of higher complexity.

In addition, the present study gives some indication of the aspect of complexity that makes implicit learning of artificial grammars increasingly difficult. Namely, the regression analyses indicated that topological entropy was the best predictor of performance. As illustrated by its high correlation with the number of elements required to predict the next (see Table 1 on page 13) and as explained in Appendix A, topological entropy is sensitive to the presence of higher-order dependencies. The results of the present study therefore suggest that implicit learning of artificial grammars is hampered most when letter chunks of increasing size have to be taken into account to determine whether or not the exemplars are grammatical. This situation arises when a link is added to the grammar that generates a letter chunk that could already be generated by an existing link. The number of associations that has to be acquired seems to be a less influential aspect of the complexity of artificial grammars.

The effect of complexity on the type of knowledge acquired in AGL

The second aim of this study was to investigate whether complexity influences the type of knowledge participants acquire from studying letter strings. The results of previous studies suggested that higher-order dependencies may be more difficult to learn implicitly than first-order dependencies (Gomez, 1997; Meulemans & Van der Linden, 1997). We therefore predicted that, if implicit learning would be hampered by increasing complexity, second-order dependencies would be particularly affected. The analyses showed that, for implicit learning, the strongest negative effect of increasing complexity was indeed observed for test exemplars containing several violations of second-order dependencies. Test exemplars that only contained first-order violations were recognized as ungrammatical irrespective of the complexity of the grammar. The results confirm that participants who engage in implicit learning more easily acquire

knowledge of first-order dependencies than of second-order dependencies. In addition, these results are in line with the suggestion that the occurrence of higher-order dependencies makes implicit artificial grammar learning increasingly difficult.

After explicit learning, participants were able to recognize the grammaticality of grammatical test exemplars and the ungrammaticality of test exemplars in which only first-order dependencies were violated. Surprisingly, performance on the latter type of exemplar was enhanced by increasing complexity. This may be due to participants' strategies in explicit learning. On the one hand, participants may restrict themselves to looking for simple first-order dependencies when they are required to deal with high complexity. Grammars of lesser complexity, on the other hand, may invite participants to try to discover higher-order dependencies. So, although explicit learning of artificial grammars produces a constant level of performance, this may be based on different types of knowledge for grammars of varying complexity.

Qualifying the effects of complexity

The present experiment demonstrated that implicit learning declined as the complexity of artificial grammars increased, particularly with regard to second-order dependencies. One might take this as support for claims as those of Gold (1967) and Pinker (1989) that implicit learning cannot underlie the acquisition of natural grammars, which are arguably far more complex than the finite state grammars used here (Chomsky, 1957). However, implicit learning may be enhanced by some characteristics of the natural situation. The acquisition of natural grammars has been proposed to be facilitated by initially focusing on fragments rather than on complete sentences (Elman, 1993; Kersten & Earles, 2001; Newport, 1990) and by semantic constraints (Rohde & Plaut, 1999). Nagata (1977) showed that the presence of a semantic reference field enabled participants to acquire a more complex artificial language than they could on the basis of sentences alone. It would be interesting to examine whether these factors could eliminate the negative influence of complexity on implicit learning of the artificial grammars used in the present study.

In addition, the scope of the finding that implicit learning of second-order dependencies is hampered by increasing complexity should be further investigated. Certain instructions (Whittlesea & Dorken, 1993) or more extended exposure to grammatical exemplars (Meulemans & Van der Linden, 1997) in the induction phase may enable participants to implicitly acquire second-order dependencies and to

maintain a relatively high performance on a grammaticality judgment task in spite of increasing complexity of the grammar.

The present study indicates that, with standard instructions and limited exposure to grammatical exemplars in the induction phase, the scope of implicit learning is limited. Although leading to better performance than explicit learning in the tested range, implicit learning was negatively affected by increasing complexity of the grammars. In addition, our results suggest that complexity affects the type of knowledge acquired from implicit learning. Whereas second-order dependencies may be acquired for relatively simple grammars, implicit learning of more complex grammars is limited to first-order dependencies. These results are at odds with theories of implicit learning that involve automatic abstraction of any structure present in the environment, regardless of its complexity.

