

Structural selection in implicit learning of artificial grammars

Esther van den Bos · Fenna H. Poletiek

Received: 5 June 2008 / Accepted: 5 January 2009 / Published online: 12 February 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract In the contextual cueing paradigm, Endo and Takeda (in *Percept Psychophys* 66:293–302, 2004) provided evidence that implicit learning involves selection of the aspect of a structure that is most useful to one's task. The present study attempted to replicate this finding in artificial grammar learning to investigate whether or not implicit learning commonly involves such a selection. Participants in Experiment 1 were presented with an induction task that could be facilitated by several characteristics of the exemplars. For some participants, those characteristics included a perfectly predictive feature. The results suggested that the aspect of the structure that was most useful to the induction task was selected and learned implicitly. Experiment 2 provided evidence that, although salience affected participants' awareness of the perfectly predictive feature, selection for implicit learning was mainly based on usefulness.

Introduction

Implicit learning is often defined as a process that occurs without the intention to learn, which results in knowledge that is not completely accessible to consciousness (e.g., Mathews et al., 1989; Reber, 1993; Seger, 1994). It is thought to underlie the acquisition of complex patterns, such as motor skills,

social rules (e.g., Seger, 1994), and the grammars of natural languages (Reber, 1976). Originally, implicit learning was conceptualized as an unintentional, ineluctable, and inflexible process. Reber (1993) proposed that implicit learning automatically abstracts knowledge of covariation patterns from the environment. In addition, it would be an evolutionarily old process; robust with respect to disorders and aging, virtually invariant between individual humans and shared by other species. The invariance of implicit learning was also stressed by Hayes and Broadbent (1988), who proposed that the distinguishing factor between implicit and explicit learning is the selectivity of the processes. Explicit learning would involve active selection of a small amount of relevant information, whereas implicit learning would unselectively store the frequency of co-occurrence of all elements present.

This passive view of implicit learning was opposed by Whittlesea et al. (Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998), who argued that implicit learning does not unselectively capture any structure present in the environment. They demonstrated that the kind of knowledge acquired in implicit learning experiments could be modified by the task participants perform on the stimuli (Whittlesea & Dorken, 1993) and by accidental characteristics of the stimuli (e.g., familiarity, salience; Whittlesea & Wright, 1997) and the context (e.g., spatial organization; Wright & Whittlesea, 1998). According to their episodic processing account, sensitivity to a structure (at test) is due to accidental overlap in information processing with earlier (learning) situations. What is learned can vary widely; structure has no special status in the selection process. Therefore, Whittlesea et al. see no need to invoke an implicit structure learning mechanism (Wright & Whittlesea, 1998). In this account, the knowledge acquired in any situation depends on what is attended and attention is treated as essentially unpredictable.

Electronic supplementary material Supplementary material is available in the online version of this article at doi:10.1007/s00426-009-0227-1 and is accessible for authorized users.

E. van den Bos (✉) · F. H. Poletiek
Leiden University, Cognitive Psychology,
P.O. Box 9555, 2300 RB Leiden, The Netherlands
e-mail: bosejvanden@fsw.leidenuniv.nl

F. H. Poletiek
e-mail: poletiek@fsw.leidenuniv.nl

The present study will explore a middle course between the view that implicit learning is unselective and the view that implicit learning is unpredictable, by proposing that people implicitly learn the aspect of a structure that is most useful to the task they perform while they encounter the structure. This view takes into account evidence from visual search tasks that the relationship between attention and learning is bi-directional (Chun & Jiang, 1998; Lambert, 2003). As proposed by Chun and Jiang (1998), people learn to focus attention on certain aspects of a structure when it is presented repeatedly (although this default can be overcome by factors such as those identified by Whittlesea & Wright, 1997). The deployment of attention, in turn, affects the way the stimuli are processed and, as pointed out by Whittlesea et al., thereby determines the contents of people's knowledge.

Chun and Jiang (1998) demonstrated that structure learning can guide attention by presenting participants in a contextual cueing experiment with search displays containing one target and several distracters. Half of the displays had configurations of targets and distracters that were repeated during the experiment. For these configurations, participants became increasingly faster at locating the targets, because they learned where to attend. Learning was considered implicit, because participants were not informed about the repeated configurations prior to the experiment and were unable to distinguish between old and new configurations on a subsequent recognition test (Chun & Jiang, 1998).

Endo and Takeda (2004) went on to investigate which aspect of the structure would guide attention. In their experiments, the location of the target could be predicted by both the spatial configuration and the identity of the distracters. When both were perfectly predictive, attention was guided by the predictor that was easiest to associate with the target location and only distracter configuration was learned. However, when each relationship predicted the target location on half of the trials, both were learned. When distracter identity was perfectly predictive and distracter configuration was made uninformative, by varying it randomly or associating it with target identity rather than location, only the identity predictor was learned. Together these experiments suggest that, by default, the aspect of a structure that is most useful to one's task guides attention and is implicitly selected for learning. This demonstration of implicit selection of useful structure is interesting, because both the basis of selection and its implicitness are controversial in the light of findings from other implicit learning paradigms. These findings will be discussed below.

Selection of useful information

On one hand, the finding that an aspect of a structure was implicitly selected for learning on the basis of its usefulness

to the visual search task (Endo & Takeda, 2004) may be expected in other tasks as well. One might argue that knowledge of the structure is useful to the participants' task in most demonstrations of implicit learning. For example, in serial reaction time (SRT) tasks (Nissen & Bullemer, 1987) participants have to react as quickly and accurately as possible to a stimulus appearing at some location on a computer screen by pressing the corresponding key on the keyboard. Knowing the sequence of locations in which the stimulus appears allows for faster responding. In the Artificial Grammar Learning (AGL) paradigm (Reber, 1967), participants who memorize letter strings without knowing that they correspond to an artificial grammar are subsequently able to discriminate between new grammatical and ungrammatical exemplars. Reber (1967) demonstrated that participants who memorized grammatical exemplars needed fewer learning trials to achieve accurate reproduction than participants who memorized stimuli that were randomly composed of the same letters. So, memorizing the individual exemplars is facilitated by acquiring knowledge of the underlying structure. This suggests that implicit learning typically occurs for structures that are useful to one's task.

On other hand, several SRT-experiments have demonstrated implicit learning of the sequence of locations in the presence of a more reliable predictor of where the next stimulus would appear. Participants were shown to learn the sequence when the location of the next stimulus could be perfectly predicted from the identity of the present stimulus (Jiménez & Méndez, 1999, 2001) and when the next location was indicated by an explicit cue (Cleeremans, 1997). This suggests that selection of the aspect of a structure that is most useful to the task one performs while encountering the structure may not be a general finding in implicit learning.

The discrepancy between the results of Endo's and Takeda's (2004) contextual cueing experiments and the findings in the SRT-paradigm (Cleeremans, 1997; Jiménez & Méndez, 1999, 2001) may reflect different learning mechanisms. There is some evidence that different neural substrates underlie the formation of the spatial associations acquired in contextual cueing experiments and the spatio-temporal associations acquired in SRT-tasks (Howard, Howard, Dennis, Yankovich, & Vaidya, 2004). Interestingly, Dominey (2003) proposed a similar distinction between a mechanism for spatiotemporal structure learning and a mechanism underlying AGL. As pointed out by Seger (1994), implicit learning of a spatiotemporal pattern in the SRT-task involves a mechanism for the planning of responses, whereas implicit learning of a visuospatial pattern, as in contextual cueing and artificial grammar learning, is likely to involve a different mechanism for evaluating the fluency with which the pattern is processed. This

suggests that selection of the aspect of a structure that is most useful to the task participants perform when they encounter the structure, observed in contextual cueing experiments (Endo & Takeda, 2004), but not in SRT-tasks (Jiménez & Méndez, 1999, 2001), may be replicated in AGL.

Implicitness

A second interesting point raised in the contextual cueing paradigm is that an aspect of a structure may be selected and learned implicitly (Chun & Jiang, 1998; Endo & Takeda, 2004). This suggestion is at odds with a previous finding that people have conscious control over selection of information. Haider and Frensch (1999) found that participants processed irrelevant information when they were instructed to perform a task as accurately as possible, but not when they were instructed to perform the task as fast as possible. They concluded that, although participants may implicitly assess whether or not information is relevant, they intentionally decide whether or not to process it.

Moreover, learning without awareness itself is controversial. Artificial grammar learning has been shown to result in explicit knowledge of bigrams (Dulany, Carlson, & Dewey, 1984), which can be sufficient to achieve normal performance on a grammaticality judgment task (Perruchet & Pacteau, 1990). In addition, Shanks and St. John (1994) argued that early studies demonstrating knowledge without awareness had failed to detect explicit knowledge, because they used insensitive measures and focused on irrelevant information. At present, it is generally acknowledged that implicit learning produces a certain amount of explicit knowledge (Cleeremans, 1993; Reber, 1993).

However, there is some evidence that explicit knowledge is insufficient to explain performance (Mathews et al., 1989) and that it may not actually be used in making grammaticality judgments (Meulemans & Van der Linden, 2003). In addition, it has been proposed that implicit learning produces knowledge that is different from explicit knowledge in the sense that it is not accompanied by meta-knowledge (Dienes & Berry, 1997). In implicit learning, people may form representations that are not labeled as knowledge and, hence, cannot be recognized as such (Dienes & Perner, 1999). In sum, the generality of the finding that an aspect of a structure can be selected and learned implicitly (Endo & Takeda, 2004) has not yet been established.

The present study further investigated both the possibility that selection for implicit learning is based on a structure's usefulness to the task people perform when they encounter it and the degree to which this selection can be performed implicitly. To address the first question, we

explored whether structural selection of the aspect most useful to one's task could be observed in artificial grammar learning. We presented participants with exemplars that contained several useful characteristics, one of which was perfectly predictive. If participants would be biased to selectively learn the aspect of the structure that is most useful to their task in the induction phase, they would learn the perfectly predictive feature rather than the other characteristics. To address the second question of this study, we investigated the degree to which learning was implicit by estimating performance on a classification test in the absence of explicit knowledge.

Experiment 1

Endo and Takeda (2004) showed that implicit learning in the contextual cueing paradigm was limited to the aspect of a structure that was most useful to the search task participants performed when they encountered the structure. To explore whether or not such selectivity could be demonstrated in another implicit learning paradigm, we presented each participant with exemplars from two artificial grammars (see Fig. 1). The grammars consist of different rules governing the order of the letters in their exemplars, resulting in the occurrence of distinctive groups of two and three letters (bigrams and trigrams). Similarly, they specify different probabilities of occurrence for the letters overall, and at each position in an exemplar. Each of these characteristics could potentially facilitate the participants' task in the induction phase, which was to memorize the exemplars and the side of the screen where they appeared. A characteristic's actual usefulness, however, would depend on the proportion of exemplars it applies to, henceforth its predictive value (the number of induction phase exemplars from one grammar in which the characteristic occurs minus the number of induction phase exemplars from the other grammar in which the characteristic occurs, divided by the total number of induction phase exemplars from a grammar, 32 here).

In the experiments of Endo and Takeda (2004), selective learning of the most useful aspect of the structure was observed for perfect predictors of the target location. In our attempt to replicate their finding, we therefore included a perfectly predictive feature in one of the stimulus sets. Our main question was whether or not people are biased to selectively learn this feature at the expense of less predictive characteristics. Therefore, participants were presented with two kinds of stimuli at test. One half of these could be classified on the basis of both the feature and the less predictive characteristics, while the other half could only be classified on the basis of the less predictive characteristics. A baseline level of learning for the other characteristics was

Fig. 1 Artificial grammars used in this study. Grammars A1 and B1 generated Stimulus Set 1, with a perfectly predictive feature. Grammars A2 and B2 generated Stimulus Set 2, without a perfectly predictive feature. Grammars A3 and B3 generated Stimulus Set 3, with a non-salient feature. The grammars are based on those of Whittlesea and Dorken (1993, Experiment 1)

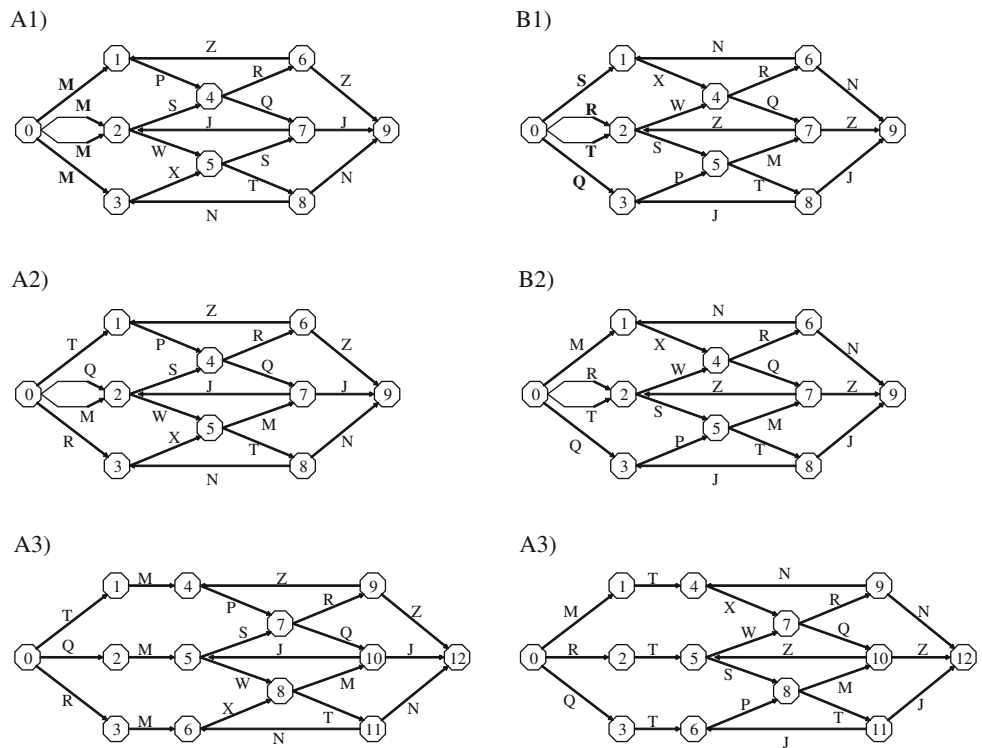


Table 1 The ten most predictive characteristics of the grammars used in Experiment 1

Feature			No feature		
Predictive value	Characteristic	Grammar	Predictive value	Characteristic	Grammar
1	Begin with M	A	0.5625	Z	B
0.5625	Z	B	0.5625	MJ	A
0.5625	MZ	B	0.5625	MZ	B
0.5625	SJ	A	0.53125	QJ	A
0.53125	QJ	A	0.53125	TN	A
0.53125	TN	A	0.46875	QZ	B
0.46875	QZ	B	0.46875	RN	B
0.46875	RN	B	0.46875	TJ	B
0.46875	TJ	B	0.46875	ZS	B
0.46875	ZS	B	0.4375	TNX	A

obtained with the stimulus set without a feature. The characteristics with the highest predictive values for each stimulus set are shown in Table 1.

Method

Participants

A total of 56 undergraduate students of Leiden University (17 males, 39 females, 17–29 years of age; $M = 20.37$, $SD = 2.86$) participated in this experiment. They signed up for either a 30-min session or a 15-min session. Students who signed up for a 30-min session were randomly

assigned to two experimental groups and students who signed up for a 15-min session were randomly assigned to two control groups, so that each group contained 14 participants. The students received either course credits or money for their participation; experimental participants were paid € 4.50 and control participants were paid € 3.

Design

Three independent variables were manipulated in the experiment. Firstly, the participants were divided into two experimental groups and two control groups. The experimental groups were presented with an induction phase, in which they had to memorize exemplars from two artificial

grammars together with the side of the screen where they appeared. In addition, they had to type in the stimuli in order to guarantee that they would attend to the letters of all exemplars. For the control groups there was no induction phase.

Secondly, the stimulus set was varied between participants. Participants in one experimental group and one control group worked with materials with a perfectly predictive feature; participants in the other experimental group and the other control group worked with materials without a perfectly predictive feature. For each participant in the (experimental) memorize groups, exemplars generated by one grammar were always presented on one side of the screen during the induction task, while exemplars generated by the other grammar were always presented on the other side. The order of presentation of the exemplars was randomized. The side of the screen (left or right) with which each grammar was associated was balanced over all participants.

Thirdly, the type of exemplar was varied within subjects in the test phase. All participants classified both complete exemplars and fragments, presented in random order. Knowledge of the less predictive characteristics of each grammar would allow for accurate performance on both complete exemplars and fragments, while knowledge of the perfectly predictive feature would allow for accurate performance on complete exemplars, but not on fragments.

The main dependent variable was the proportion of exemplars correctly classified as belonging to the side of the screen associated with their grammar. Participants in the memorize groups knew from the induction phase which side of the screen each grammar was associated with. Participants in the control groups, however, did not know which side each grammar was assigned to. If, as expected, control participants were unable to distinguish between exemplars from the two grammars, they would perform at chance level irrespective of the mapping they had been randomly assigned to. However, if they consistently classified exemplars from one grammar to one side and exemplars from the other grammar to the other side, their accuracy could be either high or low, depending on the mapping they were randomly assigned to. To check whether the accuracy results were biased by the control group's mapping, consistency was included as a second dependent variable. Consistency was defined as the difference between the number of exemplars from grammar A classified as belonging to one side and the number of exemplars from grammar B also classified as belonging to that side.

Materials

The stimuli were generated by the four finite state grammars shown in Fig. 1. The grammars were implemented in a computer program, which generated 56 different exem-

plars for each grammar, consisting of either seven or ten letters. From these exemplars two sets of stimuli were created. Set 1 consisted of the exemplars generated by A1 and B1; Set 2 consisted of the exemplars generated by A2 and B2 (see Appendix A of Electronic supplementary material). Exemplars generated by A1 always started with an M, while exemplars generated by B1 never started with an M. So, for participants who memorized exemplars from Set 1 and the side of the screen where they appeared, their task could be facilitated by this perfectly predictive feature as well as by letters occurring at other positions, bigrams, and trigrams characteristic of each grammar. Since the feature was invariant, while none of the other characteristics occurred in each exemplar from a grammar, the feature has a higher predictive value, indicating that it is more useful to the task in the induction phase than the other characteristics. In Set 2, such a feature was not available and the memorize task could only be facilitated by learning the less predictive characteristics.

Each stimulus set was divided into 64 induction stimuli and 48 test stimuli, so that both groups consisted of an equal number of exemplars from grammars A and B. One half of the test stimuli, balanced for grammar and length, were presented as "complete exemplars". The other half were presented as "fragments", created by replacing the first letter of the exemplar by an underscore. The stimuli for practice in the test phase consisted of five additional exemplars for each stimulus set: one complete exemplar and one fragment generated by grammar A and two complete exemplars and one fragment generated by grammar B. The stimuli for practice in the induction phase consisted of numbers and were unrelated to the grammars.

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 and by writing their answer to an open question on a sheet of paper.

Procedure

The participants were tested individually in a dimly lit test booth. At the beginning of the experiment, the participants in the memorize groups were told that it would consist of two parts. They were informed that they would first be presented with two groups of exemplars, a left group and a right group, and that they would have to memorize each exemplar together with the side of the screen where it appeared. Subsequently, there were five practice trials. Participants were notified when the experimental trials began. Each trial started with a fixation cross appearing on the left or on the right of the screen. After 1 s the cross was replaced by an exemplar, centered at the fixation point.

Participants performed the memorize instruction and typed in the letters. When they pressed the last button, their input was displayed on the screen underneath the original exemplar and a reminder of the instruction appeared. After 2 s the screen turned blank for 1 s and then the next trial began.

Participants in the control groups only performed the second part of the experiment, which was a classification task. At the beginning of this task, all participants were informed that they would be presented with both complete exemplars and fragments of exemplars. They were told that the stimuli would appear in the middle of the screen and that half belonged to the left and half belonged to the right group. They were required to indicate for each exemplar to which group it belonged by pressing either the left or the right arrow on the keyboard. They received 5 practice trials, followed by 48 experimental trials. Each trial began with a fixation cross appearing in the middle of the screen. After 1 s the cross was replaced by an exemplar centered at the fixation point. The exemplar remained on the screen until the participant pressed one of the arrows. Then there was a blank screen for 1 s before the next trial started.

After the participants had completed all trials, they were asked to write down whether they had noticed any differences between the left group and the right group and, if they had, what these differences were. Finally, the participants were thanked for their participation. The experiment took about 25 min for the memorize groups and 15 min for the control groups.

Analyses for type of knowledge

The question of whether or not selective learning could occur without leading to awareness was addressed by estimating performance on the classification test in the absence of explicit knowledge. Shanks and St. John (1994) have pointed out that deducting explicit knowledge from a combined measure of implicit and explicit knowledge only results in a valid estimate of implicit knowledge when the test of explicit knowledge taps the information that participants actually use and is sensitive to all relevant knowledge. In the present study, the measure of explicit knowledge was based on the responses to the open question. Verbal reports are a traditional measure of explicit knowledge, but their sensitivity is limited. Therefore, if participants based their judgments on knowledge of many bigrams and trigrams, their explicit knowledge of these characteristics may be underestimated. Knowledge of a single perfectly predictive feature, however, would be easy to verbalize and should be revealed by this question.

The answers to the question were scored following a procedure developed by Dienes, Broadbent, and Berry (1991). Firstly, the criteria provided by the participant were applied to the test stimuli in an attempt to determine for

each exemplar whether it would be classified correctly or incorrectly. Secondly, exemplars that could not be classified, since none of the participant's rules applied to it, were assumed to be guessed correctly in 50% of the cases. Therefore, the verbal report score was defined as the number of exemplars that would be classified correctly plus half the number of unclassifiable exemplars. For example, any mention of the feature would lead to correct classification of the 24 complete exemplars and correct guesses for 12 of the fragments, amounting to a verbal report score of 36. If participants did not have any relevant explicit knowledge, their verbal report score would be 24.

Subsequently, a regression analysis was performed with verbal report score as the independent variable and proportion correct on the actual classification test as the dependent variable. The resulting regression equation was used to derive a reliable estimate of the proportion of correct classifications associated with a verbal report score of 24 (no explicit knowledge). If the predicted proportion of correct classifications was significantly above chance, this indicated implicit knowledge.

Results

Induction phase performance

Memorizing each exemplar together with the side of the screen where it appeared and typing it in took an average 7.84 s (SD = 1.31) for the participants who memorized stimuli with a perfectly predictive feature. The mean proportion of exemplars typed in correctly was 0.895 (SD = 0.064) for this group. Participants who memorized stimuli without a perfectly predictive feature took on average 7.86 s (SD = 2.89) per exemplar. The mean proportion of exemplars they typed in correctly was 0.913 (SD = 0.066). There were no significant differences in exposure time ($t(26) < 1$, 95% CI = -1.77 to 1.72) or accuracy ($t(26) < 1$, 95% CI = -0.068 to 0.033) for the two stimulus sets.

Test phase performance

The mean proportions of correct classifications for each group, stimulus set, and type of exemplar are shown in Table 2. The main analysis was a 2-between, 1-within mixed model analysis of variance (ANOVA) on the proportion of correct classifications; with group (memorize, control) and stimulus set (with feature, without feature) as between-subjects variables and type of exemplar (complete exemplar, fragment) as within-subjects variable. The ANOVA showed significant main effects of group ($F(1,52) = 40.217$, MSE = 0.025, $P < 0.001$) and type of

Table 2 Mean proportions of correct classifications and standard deviations for each type of exemplar by stimulus set and group for Experiment 1 and Experiment 2

Group	Feature		No feature	
	Complete	Fragment	Complete	Fragment
Experiment 1				
Memorize	0.896 (0.148)	0.527 (0.126)	0.670 (0.157)	0.682 (0.149)
Control	0.521 (0.174)	0.503 (0.068)	0.494 (0.107)	0.500 (0.080)
Experiment 2				
Memorize	0.772 (0.178)	0.621 (0.169)		
Control	0.506 (0.081)	0.503 (0.040)		

Standard deviations are in parentheses

exemplar ($F(1,52) = 24.916$, $MSE = 0.010$, $P < 0.001$), significant two-way interactions of group and type of exemplar ($F(1,52) = 21,921$, $MSE = 0.010$, $P < 0.001$) and of stimulus set and type of exemplar ($F(1,52) = 30.049$, $MSE = 0.010$, $P < 0.001$) and a significant three-way interaction of group, stimulus set and type of exemplar ($F(1,52) = 23.395$, $MSE = 0.010$, $P < 0.001$). Because main effects and two-way interactions are qualified by this three-way interaction, they will not be interpreted.

The contribution of stimulus set to the three-way interaction is examined first, because participants were expected to learn different characteristics of the two sets of stimuli. For the stimulus set with a feature, there was a significant interaction between group and type of exemplar ($F(1,26) = 33.905$, $MSE = 0.013$, $P < 0.001$). Independent samples t tests showed that more complete exemplars were classified correctly by the memorize group than by the control group ($t(26) = 6.139$, $P < 0.001$). For fragments, however, there was no significant difference ($t(26) < 1$, 95% CI = -0.055 to 0.102). For the stimulus set without a feature, only the main effect of group was significant ($F(1,26) = 17.137$, $MSE = 0.026$, $P < 0.001$). The proportion of correct classifications was higher for the memorize group ($M = 0.676$, $SD = 0.143$) than for the control group ($M = 0.497$, $SD = 0.075$).

The ANOVA on the consistency scores showed the same pattern of results (see Table 3), indicating that the accuracy data were not biased by the way the grammars were assigned to the sides of the screen for the control group. In addition, this analysis indicated that participants who memorized exemplars containing a perfectly predictive feature did not learn any other characteristics of the grammars. If they had learned other characteristics, but failed to associate those with a side of the screen, they would have been more consistent than control participants in grouping fragments from the same grammar together. In summary, the results indicate that participants who memorized exemplars with a perfectly predictive feature were biased to learn only

Table 3 Consistency analyses

Effect	Experiment	
	1	2
Stimulus set \times Group \times Type of exemplar	$F(1,52) = 19.496^{***}$	
Without feature		
Group	$F(1,26) = 12.519^{**}$	
With feature		
Group \times Type of exemplar	$F(1,26) = 19.866^{***}$	$F(1,26) = 4.417^*$
Complete exemplars		
Group	$t(26) = 5.315^{***}$	$t(16.0) = 5.063^{***}$
Fragments		
Group	$t(26) = 1.604$	$t(14.6) = 3.287^{**}$

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

this characteristic, whereas participants who memorized exemplars without such a feature learned the less predictive characteristics of each grammar.

Type of knowledge

Participants who memorized exemplars from the stimulus set with a perfectly predictive feature reported 16 characteristic letters, 27 letters at specific positions, 3 bigrams and 6 global characteristics (see Table 1 in Appendix B of Electronic supplementary material). Their verbal report scores over all exemplars ranged from 23.5 to 39.5 out of 48 ($M = 32.9$, $SD = 5.5$). Because the performance analysis indicated that these participants had knowledge relevant to complete exemplars, but not to fragments, the present analysis was restricted to verbal report scores over complete exemplars. Those scores ranged from 11.5 to 24 out of 24 ($M = 20.4$, $SD = 5.2$). Nine participants had complete

explicit knowledge of the feature. The regression analysis showed that the verbal report score was a significant predictor of the proportion of complete exemplars classified correctly ($F(1,12) = 18.025$, $P = 0.001$). The predicted proportion of correct classifications for a verbal report score of 12, corresponding to no explicit knowledge, was 0.708, which is significantly above chance (95% CI = 0.597–0.820).

Participants who memorized exemplars from the stimulus set without a feature reported 15 letters, 17 bigrams, 16 trigrams, one group of 4 letters and 3 global characteristics (see Table 1 in Appendix B of Electronic supplementary material). Their verbal report scores over all exemplars ranged from 22 to 41.5 out of 48 ($M = 30.6$, $SD = 7.3$). The regression analysis showed that the verbal report score was a significant predictor of the proportion of correct classifications ($F(1,12) = 72.163$, $P < 0.001$). The predicted proportion of correct classifications for a verbal report score of 24, corresponding to no explicit knowledge, was 0.555, which is significantly above chance (95% CI = 0.510–0.600).

Discussion

We investigated the hypothesis that people are biased to selectively learn only the aspect of a structure that is most useful to the task they perform when they encounter the structure. The results of Experiment 1 showed that, after memorizing exemplars from two artificial grammars that could be distinguished on the basis of several characteristics, including one perfectly predictive feature, participants were able to classify complete exemplars, while they were unable to classify exemplars from which the feature had been removed. This indicates that they had learned the perfectly predictive feature, but not the other characteristics of each grammar. In contrast, participants who could only use the less predictive characteristics to facilitate their task were significantly better than the control group for both types of exemplars. Their overall performance of 67% correct was similar to the 66% correct found by Whittlesea and Dorken (1993, Experiment 1: unambiguous items), who also asked participants to classify exemplars to one of two grammars. In short, the results were in accordance with the hypothesis. When several aspects of the structure were useful to the induction task, only the most useful aspect was learned. However, the other aspects could be learned if there was no more useful alternative.

In addition, the aspect most useful to the task in the induction phase seemed to be selected and learned implicitly. For participants who memorized exemplars from the stimulus set without a feature and the side of the screen where they appeared, the regression analysis indicated a small, but significant, ability to classify new exemplars in

the absence of explicit knowledge. Although this finding should be interpreted cautiously, because the verbal reports may have been insufficiently sensitive, it suggests that participants had explicit knowledge related to some letters at certain positions, bigrams or trigrams and implicit knowledge related to others. Of the participants who memorized exemplars containing a perfectly predictive feature together with the side of the screen where they appeared, some explicitly reported the feature. Nevertheless, the regression analysis showed above chance performance on complete exemplars in the absence of explicit knowledge, probably indicating that the remaining participants had learned it implicitly. As it seems unlikely that participants who based their classifications on a single feature would fail to report it on the questionnaire, this suggests that very simple information that is useful to one's task can be learned selectively without reaching awareness.

The results of the present experiment could be taken as evidence that the finding of implicit selection and learning of useful information (Endo & Takeda, 2004) generalizes from the contextual cueing paradigm to artificial grammar learning. However, it could be argued that the feature was not only the most useful aspect of the structure in the present experiment, but also the most salient. Whittlesea and Wright (1997) have shown that people may fail to learn aspects of a structure when their attention is captured by a salient feature. The findings of Endo and Takeda (2004) suggest that usefulness to one's current task may independently draw attention to an aspect of a structure. We tentatively suggest that using any aspect of a structure to perform a task provides sufficient attention to the aspect to be learned implicitly. When an aspect of the structure is salient, however, it is likely to capture attention more fully and to be learned explicitly (Reber, Kassin, Lewis, & Cantor, 1980; though see Turner & Fischler, 1993). If the most useful aspect of the structure in Experiment 1 was also salient, the relative contributions of the two factors remain unclear. Learning may even have been restricted to the feature, because attention was fully captured by its salience, leaving the rest of the exemplar unattended.

Several aspects of the data from the present experiment make this extreme interpretation unlikely. Firstly, participants typed in 90% of the exemplars faultlessly in the induction phase, which requires at least some attention to the subsequent letters in the string. Secondly, if the salient feature were the only aspect of the structure that was attended during the induction phase, one might expect participants exposed to the feature to spend less time looking at the exemplars than participants memorizing exemplars without a salient feature. In addition, one would expect all learning of the feature to be explicit. These expectations were not supported by the data. Nevertheless, we conducted a second experiment to investigate whether selective

implicit learning of a perfectly predictive feature can occur when its salience is reduced.

Experiment 2

In this experiment, two grammars were used that contained invariant second letters instead of an invariant first letter. Frick and Lee (1995) found that 79% of participants noticed an invariant first letter in otherwise random letter sequences, whereas only 24% noticed an invariant second letter. Therefore, removing the invariant letter from the first position was expected to reduce the feature's salience, without affecting its usefulness to the task of memorizing each exemplar and the side of the screen where it appeared. The results of Endo and Takeda (2004) suggest that the feature will be implicitly selected for learning as long as it is the most useful aspect of the structure with respect to the induction task. This leads to the prediction that, as in Experiment 1, participants in Experiment 2 will be biased to learn the perfectly predictive feature rather than the other characteristics of the grammars. Alternatively, if selection of the feature was based on its salience in Experiment 1, learning of the feature would be diminished by a reduction in salience. In that case, no difference between complete exemplars and fragments would be expected in Experiment 2.

Method

Participants

There were 28 participants in this experiment (9 males, 19 females; 19–34 years, $M = 23.21$, $SD = 3.35$). All participants were students of Leiden University and none of them had participated in Experiment 1. The students could sign up for either a 30-min session or a 15-min session. The 14 students who signed up for a 30-min session were assigned to the experimental group; 14 students who signed up for a 15-min session were assigned to the control group. They received either course credits or money for their participation; experimental participants were paid € 4.50 and control participants were paid € 2.

Materials

In Experiment 2, only a stimulus set with a perfectly predictive feature was used (see Appendix A of Electronic supplementary material). The feature occurred in the second position rather than in the first. This shift in position was assumed to make the feature less salient without diminishing its predictive value in memorizing the exemplars and the side of the screen where they appeared. (See Table 4 for

Table 4 The 12 most predictive characteristics of the grammars used in Experiment 2

Predictive value	Characteristic	Grammar
1	M 2nd	A
1	T 2nd	B
0.5	QM	A
0.5	RT	B
0.46875	ZS	B
0.4375	TNX	A
0.40625	JW	A
0.40625	MZ	B
0.40625	ST	B
0.40625	WT	A
0.40625	XM	A
0.40625	Z 5th	B

the characteristics with the highest predictive values in this stimulus set). To keep the length of the exemplars from the two grammars equal and to keep the grammars as similar as possible to those used in Experiment 1, both Grammar A3 and Grammar B3 contained an invariant second letter (see Fig. 1). The exemplars varied in length from 8 to 11 rather than from 7 to 10 letters and the fragments for the test phase were created by removing the second letter instead of the first. In all other respects, the stimulus sets for Experiment 2 were created in the same way as those for Experiment 1.

Procedure and analyses

The procedure and analyses were the same as in Experiment 1.

Results

Induction phase performance

The mean proportion of exemplars typed in correctly by participants in the memorize group was 0.901 ($SD = 0.067$, 95% $CI = 0.862$ – 0.939). Memorizing each exemplar together with the side of the screen where it appeared and typing it in took on average 9.08 s ($SD = 3.89$, 95% $CI = 6.84$ – 11.33). Participants exposed to the stimulus set used in Experiment 2 did not differ in accuracy or exposure time from participants exposed to the stimulus sets used in Experiment 1 ($F(2,39) < 1$ for both measures).

Test phase performance

The mean proportions of correct classifications for each type of exemplar by group are shown in Table 2. The data

were analyzed by a 1-between, 1-within mixed model ANOVA on the proportion of correct classifications with group (memorize, control) as between-subjects variable and type of exemplar (complete exemplar, fragment) as within-subjects variable. The analysis showed significant main effects of group ($F(1,26) = 21.737$, $MSE = 0.024$, $P < 0.001$) and type of exemplar ($F(1,26) = 8.037$, $MSE = 0.010$, $P = 0.009$), which were qualified by a significant interaction between group and type of exemplar ($F(1,26) = 7.453$, $MSE = 0.010$, $P = 0.011$). As in Experiment 1, the proportion of correct classifications was higher for complete exemplars than for fragments ($t(13) = 3.002$, $P = 0.010$) in the memorize group, while there was no difference ($t(13) < 1$, 95% CI = -0.041 to 0.047) between complete exemplars and fragments in the control group. The proportion of correct classifications was higher in the memorize group than in the control group for complete exemplars ($t(18.2) = 5.103$, $P < 0.001$) and, in contrast to Experiment 1, also for fragments ($t(14.4) = 2.541$, $P = 0.023$).

The consistency analyses (see Table 3) showed a pattern of results similar to that of the accuracy analyses. Although the control group classified complete exemplars more consistently than fragments ($t(13) = 3.079$, $P = 0.009$), consistency for either type of exemplar was lower than in the memorize group.

Type of knowledge

Participants in the memorize group reported 9 characteristic letters, 13 letters at specific positions, 14 bigrams, 26 trigrams, one group of 4 letters, one group of 5 letters and 2 global characteristics (see Table 2 in Appendix B of Electronic supplementary material). The verbal report scores over all items ranged from 22.5 to 43.5 out of 48 ($M = 32$, $SD = 6.7$). The performance analysis indicated that, although participants could classify both complete exemplars and fragments, they were better at classifying complete exemplars. Therefore, the degree of implicit knowledge used to classify the two types of exemplars was estimated separately. The verbal report score for complete exemplars ranged from 10 to 24 out of 24 ($M = 17.5$, $SD = 5.1$). Three participants had complete explicit knowledge of the feature, which is a significantly lower number than in Experiment 1 ($\chi^2(1) = 5.25$, $P < 0.05$), in which nine participants had complete explicit knowledge of the feature. The regression analysis showed that the verbal report score was a significant predictor of the proportion of complete exemplars classified correctly ($F(1,12) = 11.536$, $P = 0.005$). The predicted proportion of correct classifications for a verbal report score of 12, corresponding to no explicit knowledge, was 0.639, which is significantly above chance (95% CI = 0.524 – 0.754).

The verbal report score for fragments ranged from 12 to 22 out of 24 ($M = 14.5$, $SD = 3.2$). The regression analysis showed that the verbal report score was a significant predictor of the proportion of fragments classified correctly ($F(1,12) = 6.251$, $P = 0.028$). The predicted proportion of correct classifications for a verbal report score of 12, corresponding to no explicit knowledge, was 0.541, which is not significantly above chance (95% CI = 0.433 – 0.649).

Discussion

Experiment 1 showed that, by memorizing exemplars from two artificial grammars that could be distinguished on the basis of a perfectly predictive feature and several less predictive characteristics, participants implicitly learned the feature, but not the other characteristics. Experiment 2 investigated whether selective implicit learning of the feature had been due to its salience or to its predictive value to the task of memorizing each exemplar together with the side of the screen where it appeared. We attempted to reduce the salience of the perfectly predictive feature compared to Experiment 1 by using invariant second letters instead of an invariant first letter. This led to a significant reduction in the number of participants who became aware of the feature. The 21% of aware participants in Experiment 2 is comparable to the 24% of participants who noticed an invariant second letter in the study of Frick and Lee (1995). This indicates that, although invariant second letters occurred in both grammars in Experiment 2, our manipulation to reduce the salience of the perfectly predictive feature was successful. Thus, the salience hypothesis predicted no difference between complete exemplars and fragments. The usefulness hypothesis, in contrast, predicted that participants would be able to classify complete exemplars, but not fragments from which the feature had been removed.

The performance analysis showed an intermediate outcome: although participants in the memorize group classified more complete exemplars correctly than fragments, they did better on the fragments than a control group. Subsequent analyses suggested that participants in the memorize group made use of explicit knowledge to classify the fragments. Table B2 shows that they tended to report groups of letters containing the feature rather than the feature itself. These groups of letters may have been acquired, because they also had a high predictive value. It can be seen from Table 4 that bigrams containing the feature had the highest predictive values after the feature itself. Further research will be needed to investigate the effects of overlap of useful characteristics. With regard to the present experiment, the finding that participants learned groups of letters containing the perfectly predictive feature explains why

they showed some ability to classify fragments as well.¹ Although participants could not apply their knowledge of initial trigrams containing the perfectly predictive feature as straightforwardly to fragments from which the feature had been removed as to complete exemplars, they could still match the first and third letter of the characteristic trigrams to the fragments.

The regression analysis indicated that participants would be unable to classify fragments with above chance accuracy in the absence of explicit knowledge. So, as in Experiment 1, there was no evidence of implicit learning of the less predictive characteristics of each grammar, whereas the feature could be learned implicitly. Because the feature used in Experiment 2 was much less salient, this finding suggests that the feature received selective attention because of its predictive value to the task of memorizing each exemplar and the side of the screen where it appeared. In conclusion, the finding from the contextual cueing paradigm that people are biased to selectively learn the aspect of a structure that is most useful to the task they perform when they encounter the structure (Endo & Takeda, 2004) seems to generalize to implicit learning of artificial grammars.

General discussion

Implicit learning was initially characterized as a process that automatically and unselectively captures any regularity present in the environment (Hayes & Broadbent, 1988; Mathews et al., 1989; Reber, 1993). This view was opposed by Whittlesea et al. (Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998), who demonstrated that implicit learning is selective. In their episodic processing account, what is learned is determined by what is attended. Structure learning is therefore not guaranteed. Recently, however, it has been argued that regularities in the environment can guide attention and thereby affect what is learned (Chun & Jiang, 1998; Lambert, 2003). The present study further investigated the proposal by Endo and Takeda (2004) that implicit learning involves selection of the aspect of a structure that is most useful to the task one performs when one encounters the structure. In addition, we explored the possibility raised by the work of Chun and Jiang (1998) that this structural selection can be made implicitly.

Selective learning of useful characteristics

To explore whether or not selective implicit learning of the most useful aspect of a structure can be observed outside

the contextual cueing paradigm, we tried to replicate the findings of Endo and Takeda (2004) in an AGL-experiment. In Experiment 1 of the present study, the task of memorizing exemplars from two different artificial grammars together with the side of the screen where they appeared could, for some participants, be facilitated by one perfectly predictive feature and several less predictive characteristics of each grammar. Participants presented with these exemplars acquired knowledge of the perfectly predictive feature, but not of the other characteristics. Participants who did not have a perfectly predictive feature available to facilitate their task, in contrast, learned the less predictive characteristics.

These findings suggested that participants implicitly learned the aspect of the structure that was most useful to their induction task. However, in Experiment 1, the feature was not only useful, but also salient. Experiment 2 demonstrated that a non-salient feature could be selected and learned implicitly on the basis of its predictive value. Although participants in this experiment showed some ability to classify fragments, from which the feature had been removed, their ability was based on knowledge of letter groups containing the non-salient feature, rather than of characteristics unrelated to it. Moreover, this knowledge was likely to be explicit: in the absence of explicit knowledge, participants were unable to classify fragments. Only the perfectly predictive feature was learned implicitly, even though it was not salient.

The present results were obtained using two artificial grammars and the instruction to memorize both the exemplars and the side of the screen where they appeared. However, the degree to which aspects of a structure are useful to the task people perform when they encounter the structure is also expected to influence learning of a single artificial grammar under different induction tasks. In the standard AGL paradigm, for example, frequently occurring groups of letters are more useful to the task of memorizing as many exemplars from the grammar are possible than are rare groups of letters (Poletiek & Chater, 2006). Thus, the predictive value of letter groups may also explain which characteristics of the grammar are learned when participants are instructed to memorize exemplars from one artificial grammar.

Contrary to the traditional views that characterize implicit learning as unselective (Hayes & Broadbent, 1988) and ineluctable (Reber, 1993), these results provide further evidence that implicit learning does not inflexibly acquire any structure that is present in the stimuli (c.f. Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998). They suggest at minimum that, like familiarity, salience (Whittlesea & Wright, 1997) and spatial organization (Wright & Whittlesea, 1998), usefulness of a structure to one's task may affect what people learn.

¹ Participants who did not report letter groups containing the feature did not classify more fragments correctly than the control group ($t(7.7) = 1.536, P = 0.165, 95\% \text{ CI} = -0.040 \text{ to } 0.194$).

Furthermore, findings from the contextual cueing paradigm indicate that usefulness guides attention (Chun & Jiang, 1998; Endo & Takeda, 2004), suggesting that implicit learning will reliably result in knowledge of the most useful aspect of a structure.

Unselective learning in the SRT-task?

In contrast to the suggestion that implicit learning commonly involves selection of useful information, research using the SRT-task indicated that implicit learning of the sequence of locations was not hampered by the presence of a perfectly predictive cue to the location of the next stimulus (Cleeremans, 1997; Jiménez & Méndez, 1999, 2001). As noted before, this discrepancy could be due to the involvement of different learning mechanisms in the acquisition of visuospatial and spatiotemporal associations (Dominey, 2003; Howard et al., 2004; Seger, 1994). However, there also seems to be a possibility that the seemingly redundant structure was acquired in these SRT-experiments, because it was, in fact, useful to the participants' task.

In the experiment by Jiménez and Méndez (2001), for example, participants were presented with four different stimuli occurring at four locations on the screen. Two stimuli were designated targets and the others distracters. In addition to responding to the location of the stimulus, participants had to count the number of targets they were presented with. Although the identity of the present stimulus perfectly predicted the location of the next, participants still learned the sequence of locations. This redundant structure learning, however, could be the result of the task requirement to process the stimuli as either targets or distracters rather than at the level of their unique identity. If the two targets and the two distracters were not distinguished from each other, they would only predict the next location with a validity of 50%. The simple cue would then be less predictive than the finite state grammar, which determined the next location on 80% of the trials. So, this study may be more like the contextual cueing experiment in which two aspects of the structure predicted the target location on half of the trials and both were acquired (Endo & Takeda, 2004). Further research will be needed to establish whether or not selective learning of the most useful aspect of a structure can be observed with the SRT-task.

Other questions

In addition, the scope of the effect of usefulness on AGL and contextual cueing will have to be investigated. A failure to learn other aspects of the structure has so far only been observed in the presence of a 100% predictive feature. Therefore, it remains to be seen whether the most useful

aspect of a structure is also learned at the expense of other aspects when the most useful aspect is less than perfectly predictive and when the difference in usefulness between them is smaller.

Moreover, future experiments will have to clarify how usefulness produced its effects in the present study and in the experiments of Endo and Takeda (2004). Like Wright and Whittlesea (1998), we assume that an aspect has to be attended to be learned. We propose that the perfectly predictive feature drew attention in this study, because it was useful to the task the participants performed when they encountered the structure. It is less clear, however, why participants did not learn other aspects of the structure. One possible explanation would be that the useful feature withdrew attention from those aspects. Alternatively, aspects of a structure that participants do not use while performing a task may never draw sufficient attention to be learned. In the latter case, relatively useless aspects of a structure would not be learned even if there were no highly useful alternative. Ultimately, this would imply that implicit learning does not occur if the structure is completely useless to the task the participants perform while they are exposed to it.

In a study using a Stroop paradigm, Perlman and Tzelgov (2006) provided evidence against the latter possibility. In the relevant condition of their experiments, they presented participants with color words appearing in a fixed sequence. The participants' task in the induction phase was to name the color in which the words were printed, and this varied randomly. Although paying attention to the identity of the words seems useless or even disruptive to this task, these participants were better at predicting the identity of the next word on a subsequent test than a control group. This finding suggests that the possibility that the perfectly predictive feature drew away attention from other aspects of the structure may be a better explanation for the lack of implicit learning of relatively useless characteristics in the present study.

Selective learning in a broader context

Although it is not yet clear how selective learning of the aspect that is most useful to one's current task at the expense of other aspects was achieved in the present study, it should be noted that similar findings have emerged outside the typical implicit learning paradigms. For example, it has been suggested that second language learners often fail to acquire tense markings on the verb, because the presence of temporal adverbs makes them redundant in understanding the meaning of sentences (Ellis, 2005). Similarly, according to a formal analysis of grammar induction, rules are only represented when they provide the simplest possible description of the language input that has been received (Chater & Vitányi, 2007).

Moreover, the failure to learn additional information in the presence of a useful cue is a robust finding in classical conditioning. When two conditioned stimuli reliably precede an unconditioned stimulus, the stronger conditioned stimulus is likely to overshadow the other (e.g., Mackintosh, 1971). For example, blocking experiments have shown that no association is formed between a conditioned stimulus (e.g., a light) and an unconditioned stimulus (e.g., a shock) when the unconditioned stimulus has already been associated with another conditioned stimulus (e.g., a tone; see Rescorla & Holland, 1982, for review). An eye-tracking study provided evidence that the blocking effect is due to learned attention: participants spent little time looking at the redundant cue (Kruschke, Kappenman & Hetrick, 2005). Similar studies in the AGL or contextual cueing paradigms may clarify the contribution of attentional processes to selective learning of the most useful aspect of a structure in implicit learning.

Selecting and learning implicitly

A second issue investigated by the present study is whether selective learning of the aspect of a structure that is most useful to a person's task can occur without awareness. In both experiments, participants acquired implicit as well as explicit knowledge of the difference between the two groups of exemplars. The more salient the perfectly predictive feature used in the experiment, the higher was the number of participants who became aware of it. Nevertheless, participants also acquired implicit knowledge of both the salient and the non-salient feature. This suggests that even very simple information can be selected and learned implicitly.

This suggestion is in accordance with findings by Frick and Lee (1995), who presented participants with pseudo-random letter sequences containing an invariant letter at one position. Sequences containing the invariant letter at that position were judged to be more familiar than random sequences of the same letters, even by participants who remained unaware of the invariant. The authors concluded that very simple information that would be easy to articulate could nonetheless be learned implicitly. The present experiments suggest that implicit learning of a highly useful aspect of a structure may even occur without leading to awareness when it has to be selected from among other potentially useful aspects.

In summary, the present study indicates that implicit learning of artificial grammars does not occur unselectively. Participants learned the aspect of a structure that was most useful to the task they performed when they encountered the structure. This finding is in line with the view that selective attention affects the kind of knowledge acquired in implicit learning (Whittlesea & Wright, 1997; Wright &

Whittlesea, 1998) and suggests that attention may be guided by usefulness (c.f. Endo & Takeda, 2004). By conceptually replicating the finding from the contextual cueing paradigm that an aspect of a structure can be learned selectively on the basis of its usefulness (Endo & Takeda, 2004), this AGL-study provides evidence that such selection is a common component of implicit learning. In addition, the results suggest that the aspect of a structure that is most useful to one's task may be selected and learned without reaching awareness.

Acknowledgments During part of this research Esther van den Bos was supported by a grant from the Niels Stensen Stichting. We thank Christianne Engelse for creating a pilot version of the experiment. In addition, we thank several reviewers for their comments on the earlier versions of this paper. We particularly like to thank the anonymous reviewers who suggested the quantification of usefulness, the consistency analyses and the procedure to estimate implicit knowledge.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Chater, N., & Vitányi, P. (2007). 'Ideal learning' of natural language: Positive results about learning from positive evidence. *Journal of Mathematical Psychology*, *51*, 135–163.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.
- Cleeremans, A. (1993). *Mechanisms of implicit learning. Connectionist models of sequence processing*. Cambridge: The MIT Press.
- Cleeremans, A. (1997). Sequence learning in a dual-stimulus setting. *Psychological Research*, *60*, 72–86.
- Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review*, *4*, 3–23.
- Dienes, Z., Broadbent, D., & Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 875–887.
- Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.
- Dominey, P. F. (2003). Structure and function in sequence learning. Evidence from experimental, neuropsychological and simulation studies. In L. Jimenez (Ed.), *Attention and Implicit learning. Advances in Consciousness Research*, *48*. Amsterdam: John Benjamins.
- Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305–352.
- Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception and Psychophysics*, *66*, 293–302.
- Frick, R. W., & Lee, Y. S. (1995). Implicit learning and concept learning. *The Quarterly Journal of Experimental Psychology*, *48A*, 762–782.

- Haider, H., & Frensch, P. A. (1999). Information reduction during skill acquisition: the influence of task instruction. *Journal of Experimental Psychology: Applied*, *5*, 129–151.
- Hayes, N. A., & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition*, *28*, 249–276.
- Howard, J. H., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, *18*, 124–134.
- Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 236–259.
- Jiménez, L., & Méndez, C. (2001). Implicit sequence learning with competing explicit cues. *The Quarterly Journal of Experimental Psychology*, *54A*, 345–369.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 830–845.
- Lambert, T. (2003). Visual orienting, learning and conscious awareness. In L. Jimenez (Ed.), *Attention and Implicit learning. Advances in Consciousness Research*, *48*. Amsterdam: John Benjamins.
- Mackintosh, N. J. (1971). An analysis of overshadowing and blocking. *Quarterly Journal of Experimental Psychology*, *23*, 118–125.
- 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.
- Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain and Cognition*, *52*, 250–257.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.
- Perlman, A., & Tzelgov, J. (2006). Interactions between encoding and retrieval in the domain of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 118–130.
- Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264–275.
- Poletiek, F.H., & Chater, N. (2006). Grammar induction profits from representative stimulus sampling. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, 1968–1973.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.
- Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. London: Oxford University Press.
- Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human learning and Memory*, *6*, 492–502.
- Rescorla, R. A., & Holland, P. C. (1982). Behavioral studies of associative learning in animals. *Annual Review of Psychology*, *33*, 265–308.
- 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 Sciences*, *17*, 367–447.
- Turner, C. W., & Fischler, I. S. (1993). Speeded tests of implicit knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1165–1177.
- 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., & 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.
- 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.