Processing Strategies, Fluency, and Confidence in Artificial Grammar Learning

Tobias Johansson (Tobias.Johansson@psychology.lu.se)
Department of Psychology, Lund University
Box 213, 221 00, Lund, Sweden

Abstract

According to episodic theories of implicit learning general knowledge of a domain is not directly stored in memory, but a result of the use of episodic representations. Such theories are consistent with effects of processing fluency on classification in implicit learning tasks. The experiments reported in this paper investigated whether masked priming influences classification in artificial grammar learning. The results showed no effect of masked priming under standard conditions (Experiment 1), but an effect when a response deadline was invoked (Experiment 2). In addition, the participants´ confidence judgments showed different patterns depending on the presence of a response deadline. The results are consistent with unitary episodic theories of implicit learning according to which the nominal status of a task is no guarantee for the processes it triggers.

Keywords: implicit learning; artificial grammar learning; processing fluency; confidence.

Introduction

A common theme in contemporary psychological research is *implicit cognition*, or cognition outside awareness. One aspect of implicit cognition is implicit learning, a term initially coined by Reber (1967), who found that subjects, after having been incidentally exposed to exemplars (strings of letters) generated by a complex set of rules (an artificial grammar), were able to classify previously unseen exemplars generated by the same rules above chance, even though the subjects could say very little about these rules. The experimental paradigm is commonly known as *artificial grammar learning*.

On a general level, there are two types of theories of the knowledge responsible for classification in artificial grammar learning, namely *abstractionist* (Reber, 1989) and *episodic* theories (Whittlesea & Dorken, 1993). The former suggest that classification in artificial grammar learning is based on automatically abstracted general knowledge (e.g. rules or prototypes), while the latter suggest that classification is based on episodic representations. Episodic, but not abstractionist, theories suggest that classification can be influenced by manipulations of processing fluency, especially under conditions encouraging non-analytic processing (Kinder, Shanks, Cock, & Tunney, 2003; Whittlesea & Price, 2001). This hypothesis was tested in the experiments reported in this paper. Furthermore, implicit learning is often held to result in largely implicit or unconscious knowledge. A common way of estimating the degree to which knowledge is implicit is to assess the degree to which participants have metaknowledge about their own knowledge (Dienes & Perner, 1999; Tunney; 2005). However, the experiments in this article illustrate some potential problems with using metacognitive measures in order to argue for the implicitness of knowledge representations.

Abstraction vs. Episodes

According to Reber (1967, 1989) the classification performance in artificial grammar learning can be explained by positing an implicit system that automatically abstracts the general structure of a domain. The implicit system is thus directly devoted to a process of abstraction of structure. Although such an abstractionist account can account for many of the findings in artificial grammar learning, there are a number of problems for any abstractionist view of artificial grammar learning.

First, there is no a priori need for a cognitive system to directly abstract the general structure of a domain, since the sensitivity to general structure can emerge as by-products of distributed storage of relatively raw processing experiences, i.e. episodic representations. For example, Vokey and Brooks (1992) showed that participants in artificial grammar learning were sensitive to both the grammatical status of the test items and to the item-specific similarity of test items to the training items. Orthogonal to grammaticality, the authors manipulated the test items so that half of the test items were very similar to a specific training item (a unique training item for each test item of this category) and half of the test items were dissimilar from each of the training items. The results showed that the participants were sensitive to both grammaticality and item-specific similarity. The effect of item-specific similarity is difficult to explain for an abstraction account that suggests that performance in artificial grammar learning depends only on directly abstracted general knowledge. In addition, the authors argued that the effect of grammaticality may be explained without resort to direct abstraction of general structure. Instead, the effect of grammaticality could be a result of so called “ganged retrieval” where a large number of stored training instances are activated in parallel. Importantly, the effect of grammaticality may be best described as latent implicit knowledge, rather than directly computed implicit knowledge. The emergence of knowledge of grammaticality is created as much by conditions during the test phase as the representations laid down during training. The flexibility with which people show sensitivity to general structure in different senses can often be controlled by the conditions of the test phase, suggesting that the general structure of a domain is not computed independent of constraints and intentions created by the act of probing for knowledge in a specific way (Whittlesea & Dorken, 1997).

Second, many studies have shown effects of manipulations of processing fluency, i.e. the ease with which an item is processed, on performance in a wide variety of tasks. For example, Kinder et al. (2003) showed that a perceptual clarification procedure influenced classification performance in artificial grammar learning. Since grammaticality is almost always correlated with some notion of similarity in artificial grammar learning, it can be argued that sensitivity to grammaticality is in fact only sensitivity to processing fluency produced by the similarity between grammatical test items and the training items, without any direct automatic abstraction of general structure. Importantly, effects of processing fluency are not limited to a single task. Such effects have also been found on recognition judgments (e.g. Jacoby & Whitehouse, 1989) and preference judgments (e.g. Whittlesea & Price, 2001).

According to episodic theories of implicit learning (Kinder et al., 2003; Whittlesea & Dorken, 1993) classification in artificial grammar learning might proceed as follows. During the incidental learning phase the participants process the training items in some way. The experiences of that processing are stored in memory, and there is no automatic abstraction of rules etc. During classification the net result will usually be that grammatical items are more fluently processed than ungrammatical items. The distal reason for that increased fluency (i.e. the grammar) is probably not available for the participants, so the increased fluency will feel unexpected. An unconscious inference is then made to attribute that unexpected fluency to a plausible source made salient by the task at hand, which in a classification task is the dimension of grammaticality. In effect, grammatical strings will be endorsed more often than ungrammatical strings. This use of a so called *fluency heuristic* is not the only available option. There are also many other potential heuristics that might be used. For example, the participants could scan the details of the strings and see whether any of those details trigger additional contextual detail which might give rise to recollective experience and use that as a heuristic of grammaticality. This would be an example of a so called *recollection heuristic* (Kinder et al., 2003).

Analytic vs. Non-analytic Processing

What determines whether a fluency heuristic or some other heuristic is used in a task? According to some researchers (e.g. Whittlesea & Price, 2001), the nominal status of a task (e.g. as implicit or explicit in some sense) is no guarantee for the processes it triggers. Rather, it is the combination of stimulus structure, expectations, situational constraints, and previous learning history (and probably other factors as well) that determine what processes are triggered by a task on any occasion. This perspective was nicely illustrated by Whittlesea and Price (2001).

Whittlesea and Price investigated the classic dissociation between liking and recognition in the context of the mere exposure effect. The standard result is that participants give higher liking ratings to previously seen items than to new items, but perform recognition at chance level. The standard explanation of that is that participants use the extra fluency created by previous exposure to old items as a heuristic to inform their liking ratings. However, why is it that the same heuristic is not used while performing recognition judgments, especially since it has been shown that fluency can affect recognition judgments in other contexts (e.g. Jacoby & Whitehouse, 1989)? Whittlesea and Price suggested that the answer lies in the difference between analytic and non-analytic processing strategies.

Analytic processing means processing the details of an item, while *non-analytic* processing means processing an item as a whole. Non-analytic processing is usually required if an item as a whole is to be experienced more fluently than expected. Whittlesea and Price argued that liking judgments usually trigger non-analytic processing while (difficult) recognition judgements usually trigger analytic processing. Importantly, the authors were able to reverse the dissociation by making the participants shift processing strategies. When the participants performed analytic liking judgments (by justifying their liking judgments) there was no differential preference for old over new items, and when they performed non-analytic recognition (triggered by an item pool of heterogenous categories) they performed recognition above chance level.

Kinder et al. (2003) obtained similar results as those of Whittlesea and Price in the context of artificial grammar learning, by observing a connection between non-analytic processing and a fluency effect on performance created by a perceptual clarification procedure. In addition to previous exposure and perceptual clarification rate, there are many other ways of manipulating processing fluency. One such way is the masked priming procedure used by Jacoby and Whitehouse (1989). In their experiments the participants were tested in a recognition task. On some trials of the test phase the target item was flashed briefly on the screen before it finally appeared on the screen for a recognition judgement. Presumably, this created unexpected extra fluency of processing the target item. The results showed that flashing the target item increased the probability of calling the item old. However, in a different group where the target item was flashed on the screen for a duration long enough for the item to be consciously identified the opposite result was found. The participants were now more likely to call "primed" items new, presumably because they could discount the extra fluency of processing as irrelevant to the task at hand.

The experiments reported in this paper used the same masked priming procedure as that of Jacoby and Whitehouse (1989) adopted to the context of artificial grammar learning. On half of the trials in the test phase the target item (a letter string) was flashed briefly on the screen before it appeared for classification (see Method section for details). In Experiment 1 there was no response deadline for classification, but in Experiment 2 the participants had to respond within 2 seconds on each classification trial. A response deadline was introduced as a way of encouraging non-analytic processing, since a response deadline presumably makes it harder for the participants to scan the details of the strings.

Abstractionist accounts of artificial grammar learning do not predict a fluency effect on performance. Thus, flashing the target item on the screen should not influence classification. In contrast, episodic theories, together with the assumptions of the concepts of analytic and non-analytic processing, predict that increased fluency of processing increases the probability of calling an item grammatical. Thus, episodic theories predict a fluency effect on classification in Experiment 2, where non-analytic processing is encouraged.

Metacognition and Artificial Grammar Learning

Implicit learning is often held to give rise to largely implicit knowledge, which, in turn, is often held to result in largely unconscious knowledge (Dienes & Perner, 1999; Reber, 1989). A common way to estimate the degree to which knowledge is implicit in artificial grammar learning is to assess the degree to which participants have metaknowledge about their knowledge, i.e. the degree to which participants can monitor the accuracy of their own knowledge (Allwood, Granhag, & Johansson, 2000; Tunney, 2005). The standard procedure is to ask the participants to state their confidence in having made a correct answer on each classification trial. If there is a poor relation between confidence and accuracy, then the knowledge is more likely to be implicit than explicit (Dienes & Perner, 1999). There are many ways to implement metacognitive measures of awareness in artificial grammar learning (see e.g. Allwood, Granhag, & Johansson, 2000). The experiments in this paper followed the procedure used by Tunney (2005). This procedure is further illustrated in the Results and Discussion sections below.

Some researchers have critizised the use of metacognitive measures of awareness as a way of identifying implicit knowledge (e.g. Whittlesea & Dorken, 1997). For example, metacognition is probably a largely inferential activity (Schwartz, Benjamin, & Bjork, 1997). Thus, having accurate metaknowledge about a domain does not necessarily mean that the underlying representations are explicit. It could simply mean that one is relying on cues and unconscious inferences that happen to lead to accurate metacognitive judgments.

The experiments in this paper illustrate a fundamental problem with metacognitive measures of awareness, namely that the degree of metaknowledge exhibited by participants can vary depending on the processing strategies used in the test phase (Experiment 1 vs. Experiment 2), even though the learning phase is held constant.

Experiment 1

Experiment 1 investigated the effects of masked priming on classification in artificial grammar learning, using a method similar to that of Jacoby and Whitehouse (1989). The experiment also investigated the degree to which participants have metaknowledge about their knowledge in artificial grammar learning.

Method

Thirty-two introductory psychology students took part in the experiment.

The training strings corresponded to Training List 1 (16 strings) of Vokey and Brooks (1992) and the test strings corresponded to Test List 1 and 2 (64 strings; 32 grammatical and 32 ungrammatical) of the same authors. Orthogonal to the grammaticality of the test strings, half of the test strings were also of high item-specific similarity (differing by only one letter from the most similar training string) and half of low item-specific similarity (differing by at least two letters from each of the training strings). For example, one of the training strings was MXRKVXK, which was altered into two different test strings, both of which were highly similar to it, although one was grammatical and one ungrammatical: MXRKMXK (grammatical) and MXRKRXK (ungrammatical). The length of the strings varied from three to seven letters.

During training the participants were exposed to the training strings. On each trial a string appeared on the screen for 3 s and the participants were asked to type the string back onto the screen using a keyboard after it disappeared. Each of the 16 training strings were presented in each of 3 blocks, resulting in a total of 48 training trials.

After training the participants were informed that the material was governed by complex rules and that they would now have to classify new previously unseen strings, half of which followed the rules and half of which did not. After classifying a string the participants were also asked to state their confidence regarding the correctness of the answer on a scale from 50% to 100 % certainty, where 50 meant “pure guess” and 100 meant “completely certain”.

Each trial started with a 1 s prompt on the screen that said “Be prepared!”. After that, a sequence of symbols, “#####” (the pre-mask), appeared on the screen for 500 ms, followed by a 35 ms appearance of either the target string (prime trials) that was soon to be classified on the same trial or a sequence of “+++++” (no prime trials). After that, the sequence of symbols, “#####” (the post-mask), appeared on the screen for 300 ms, followed by an empty screen for 200 ms. The target string then appeared on the screen by itself for 1 s after which a request to classify the string appeared below, with the string still on the screen. The string remained on the screen until a classification response was given, after which a confidence judgment was probed for. The number of masking symbols “#####” and “+++++” symbols (the latter only on no prime trials) were always the same as the number of letters in the target string later to appear on the screen on the same trial. The entire test set of 64 letter strings were shown twice in two different blocks, with the variables of grammaticality and item-specific similarity nested within prime and no prime trials over the two blocks. The partitioning of the test lists into prime trials and no prime trials was fixed and the same for each participant. Block order was balanced across the participants. The order of presentation within lists was random for each participant.

Results

Only the major results are reported. The results for classification and confidence are reported separately below. An α-level of .05 was adopted in all analyses.

Classification

The endorsement rates (probability of responding "grammatical") were entered into a three-way repeated measures ANOVA with the factors *Prime* (2 levels: prime vs. no prime), *Grammaticality* (2 levels: grammatical vs. ungrammatical), and *Similarity* (2 levels: high vs. low item-specific similarity).

The results showed main effects of Grammaticality, [$F(1,31) = 60.71$, $MSE = .02$, $p < .001$, $\eta^2_p = .66$], and Similarity, [$F(1,31) = 86.22$, $MSE = .97$, $p < .001$, $\eta^2_p = .74$]. Grammatical strings received a higher endorsement rate (M = .57, SE = .02) than ungrammatical strings (*M* = .43, *SE* = .02), and high similarity strings received a higher endorsement rate (*M* = .56, *SE* = .02) than low similarity strings (*M* = .44, *SE* = .02). There was no significant effect of Prime, ($F < 1$). None of the interaction effects were close to being significant (all *p*s > .15).

Confidence

The confidence judgments were entered into a four-way repeated measures ANOVA with the factors *Prime* (2 levels: prime vs. no prime), *Grammaticality* (2 levels: grammatical vs. ungrammatical), *Similarity* (2 levels: high vs. low item-specific similarity), and *Response* (2 levels: endorsement vs. rejection). One participant did not have data in each of the cells of the design and was excluded from the analysis. The mean total level of confidence was 68.90 (*SE* = 1.64).

There was a significant interaction between Similarity and Response, [$F(1,30) = 8.31$, $MSE = 17.09$, $p < .01$, $\eta^2_p = .22$], but not between Grammaticality and Response, [$F(1,30) = 3.63$, $MSE = 21.17$, $p = .07$, $\eta^2_p = .11$].

Simple main effects were conducted in order to investigate the interaction between Similarity and Response. There was an effect of Response for close items, [$F(1,30) = 14.08$, $MSE = 17.97$, $p < .01$, $\eta^2_p = .32$], but not for far items, [$F(1,30) = 2.64$, $MSE = 21.18$, $p = .12$, $\eta^2_p = .08$]. As shown in Figure 1B, this patterns was mainly driven by a significant decrease in confidence with decreasing similarity (from far to close) for rejections, [$F(1,30) = 6.75$, $MSE = 3.78$, $p < .05$, $\eta^2_p = .18$]. Thus, far strings were rejected more confidently than close strings. The decrease in confidence with decreasing similarity for endorsements was not significant, [$F(1,30) = 1.31$, $MSE = 8.74$, $p = .26$, $\eta^2_p = .04$].

Discussion

Experiment 1 replicated the basic results of Vokey and Brooks (1992), namely that there were main effects of both grammaticality and item-specific similarity on classification. There was no effect of the masked priming procedure on classification, consistent with abstractionist accounts of classification in artificial grammar learning. However, as described in the introduction, it is also possible that the potential extra fluency of processing created by the masked priming procedure is not incorporated into the classification decisions unless a non-analytic processing strategy is adopted to some extent. In order to investigate this a response deadline was introduced in the test phase of Experiment 2, since a response deadline presumably encourages a higher degree of non-analytic processing.

The participants in Experiment 1 were sensitive to both grammaticality and item-specific similarity in their classifications, but they were metacognitively sensitive only to item-specific similarity; the participants were more confident when endorsing than when rejecting items of high item-specific similarity, a difference that was not found for items of low item-specific similarity.

One way to interpret the metacognitive difference between item-specific similarity and grammaticality is that knowledge of item-specific similarity is explicit and knowledge of grammaticality is implicit. Different versions of such dual (implicit vs. explicit) knowledge bases have been offered in the literature on implicit learning (Dienes & Perner, 1999; Reber, 1989).

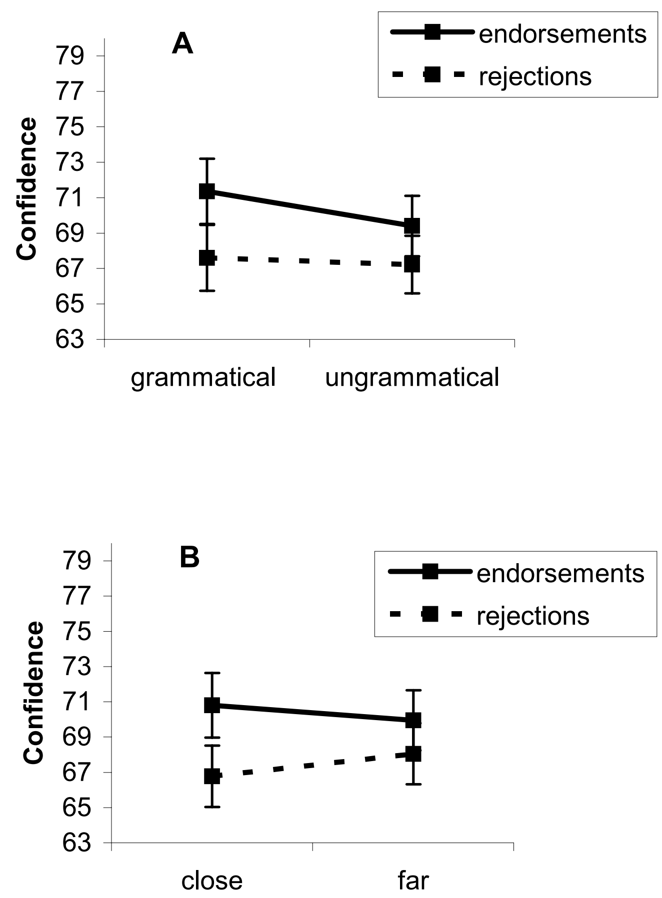


Figure 1: A: Confidence as a function of Grammaticality and Response in Experiment 1. B: Confidence as function of Similarity and Response in Experiment 1. Error bars indicate standard errors. Close and Far refer to High and Low similarity strings respectively.

However, in the material used in artificial grammar learning grammaticality is often correlated with non-specific similarity notions. That is, grammatical items are usually more similar than ungrammatical items to the training list as a whole.[1] It is possible that non-specific distributed notions of similarity are more difficult to translate into corresponding levels of confidence unless a non-analytic processing strategy is encouraged. From an episodic processing perspective, the introduction of a response deadline in Experiment 2 might then be expected to increase metacognitive sensitivity to grammaticality. However, if confidence is only dependent on whether the knowledge representations are implicit or explicit, then there should be no evidence of metaknowledge of grammaticality in Experiment 2 either.

Experiment 2

Experiment 2 was similar to Experiment 1 except that the participants had to classify each string as grammatical or ungrammatical within a response deadline of 2 s.

Method

Thirty-two introductory psychology students took part in the experiment. The material was identical to that of Experiment 1.

The procedure was identical to that of Experiment 1 in all aspects except for the classification trials during the test phase. After the postmask ("#####") a to-be-classified target string appeared on the screen. If a classification response was made within 2 seconds after the target strings appeared, then the confidence judgment window appeared on the screen and the participant was asked to make a confidence judgment without any response deadline. If a classification response was not made within 2 seconds a message appeared on the screen saying "You responded too slowly. Try to respond faster.", after which the next trial started.

Results

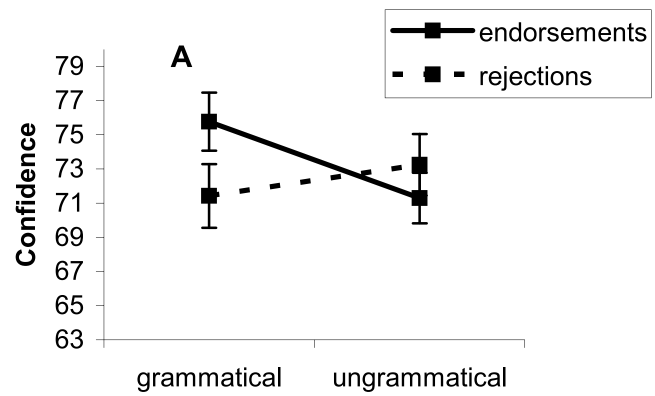
The analyses were conducted in the same way as in Experiment 1.

Classification

In contrast to Experiment 1, the ANOVA on classification judgements showed a small effect of Prime, [$F(1,31) = 5.04$, $MSE = .01$, $p < .05$, $\eta^2_p = .14$]. Target strings shown on prime trials were more likely to be endorsed ($M = .53$, $SE = .02$) than target strings on no prime trials ($M = .50$, $SE = .01$).

There were also main effects of Grammaticality, [$F(1,31) = 145.03$, $p < .001$, $MSE = .02$, $\eta^2_p = .82$], and Similarity, [$F(1,31) = 54.11$, $MSE = .01$, $p < .001$, $\eta^2_p = .64$]. As in Experiment 1, grammatical strings were endorsed more often than ungrammatical strings, and high similarity strings more often than low similarity strings.

The interaction between Grammaticality and Similarity approached significance, [$F(1,31) = 3.47$, $MSE = .01$, $p = .07$, $\eta^2_p = .10$]. The reason for this tendency was that the difference in endorsement between high and low similarity strings was numerically smaller for grammatical strings than for ungrammatical strings. None of the other interactions were close to being significant (all Fs < 1).



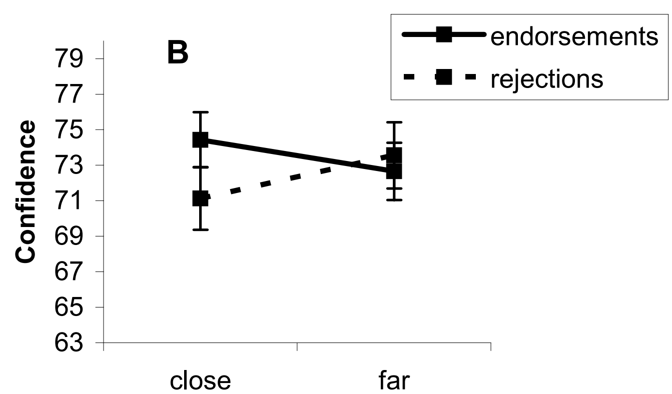


Figure 2: A: Confidence as a function of Grammaticality and Response in Experiment 2. B: Confidence as function of Similarity and Response in Experiment 2. Error bars indicate standard errors. Close and Far refer to High and Low similarity strings respectively.

[1] For example, the grammatical test items had higher *bigram chunk strength* than ungrammatical items, t(62) = 4.51, p < .01, and also higher *anchor chunk strength*, t(62) = 2.53, p < .05, but test items with high item-specific similarity did not differ from items with low item-specific similarity (both ps > .6). Both of the chunk strength indices are measures of the extent to which bigrams in a test string occurred in the set of training strings (see e.g. Tunney, 2005). Bigram chunk strength is based on all bigrams of a test string and anchor chunk strength is based on the first and last bigram of a test string. Of course, it cannot be ruled out that item-specific similarity is not also correlated with some non-specific notion of similarity. Nevertheless, item-specific similarity is certainly less correlated with non-specific notions of similarity than grammaticality is.

Confidence

The mean total level of confidence was 72.93 (SE = 1.59).

The ANOVA showed a significant interaction between Grammaticality and Response, [$F(1,31) = 36.08$, $MSE = 35.16$, $p < .001$, $\eta^2_p = .54$]. The interaction is depicted in Figure 2A. Simple main effects revealed that there was an effect of Response for grammatical strings, [$F(1,31) = 15.24$, $MSE = 19.87$, $p < .001$, $\eta^2_p = .33$], but only a marginally significant effect for ungrammatical strings, [$F(1,31) = 3.18$, $MSE = 19.08$, $p = .09$, $\eta^2_p = .09$]. Consistent with Figure 2A, this pattern was due to a significant increase in confidence with increasing grammaticality (i.e. from ungrammatical to grammatical strings) for endorsements, [$F(1,31) = 39.63$, $MSE = 8.06$, $p < .001$, $\eta^2_p = .56$], and a marginally significant decrease in confidence with increasing grammaticality for rejections, [$F(1,31) = 3.91$, $MSE = 13.67$, $p = .06$, $\eta^2_p = .11$].

There was also a significant interaction between Similarity and Response, [$F(1,31) = 19.13$, $MSE = 18.27$, $p < .001$, $\eta^2_p = .38$], as depicted in Figure 2B. Simple main effects revealed that there was an effect of Response for close strings, [$F(1,31) = 8.67$, $MSE = 20.13$, $p < .01$, $\eta^2_p = .22$], but not for far strings, ($F < 1$). As can be seen from Figure 2B, the interaction was driven by an increase in confidence with increasing similarity (i.e. from far to close strings) for endorsements, [$F(1,31) = 7.83$, $MSE = 6.44$, $p < .01$, $\eta^2_p = .20$], and a decrease in confidence with increasing similarity for rejections, [$F(1,31) = 9.30$, $MSE = 10.15$, $p < .01$, $\eta^2_p = .23$].

Discussion

In contrast to Experiment 1 the results of Experiment 2 showed a significant effect of the masked priming procedure. Flashing the target string briefly before classification made participants more likely to endorse the string as grammatical. This is difficult to explain for abstraction theories of artificial grammar learning according to which fluency of processing is not part of the relevant knowledge base of artificial grammar learning. Episodic theories, on the other hand, are more apt to handle these results. Assuming that a response deadline encourages non-analytic processing, then fluency effects on performance can be expected.

In addition, the participants in Experiment 2 were metacognitively sensitive to both grammaticality and item-specific similarity, as revealed by the interactions between grammaticality and response (endorsements vs. rejections) and item-specific similarity and response. Thus, it seems that participants can show relatively good levels of metaknowledge, as long as they engage in appropriate levels of different processing strategies. The fact that Experiment 2, but not Experiment 1, showed evidence of metaknowledge of grammaticality lends support to the idea that metaknowledge of grammaticality is enhanced by a non-analytic processing strategy, whereby non-specific notions of similarity are more appropriately incorporated into more accurate levels of confidence.

The learning phases of Experiment 1 and 2 were identical, so it is reasonable to assume that the participants developed similar representations during training. Thus, it is problematic to tie a particular level of metaknowledge to an implicit representational format of the underlying knowledge.

Taken together, the results reported in this paper lend support to episodic theories of artificial grammar learning. An important part of future research will be to further map the terrain between different fluency manipulations and different kinds of processing strategies. In addition, it would be interesting to investigate whether fluency and grammaticality effects in artificial grammar learning develop in similar ways over time in the decision process. If the grammaticality effect is based on fluency, then one would expect them to develop in a similar fashion, although with different magnitudes.

References

Allwood, C. M., Granhag, P. A., & Johansson, H. (2000). Realism in confidence judgements of performance based on implicit learning. *European Journal of Cognitive Psychology*, *12*, 165-188.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735-808.

Jacoby, L. L. & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*,126-135.

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

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855-863.

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

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential basis of metamemory. *Current Directions in Psychological Science*, *6*, 132 - 137.

Tunney, R. J. (2005). Sources of confidence in implicit cognition. *Psychonomic Bulletin and Review*, *12*, 367-373.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328-344.

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

Whittlesea, B. W. A., & Dorken, M. D. (1997). Implicit learning: Indirect, not unconscious. *Psychonomic Bulletin and Review*, *4*, 63-67.

Whittlesea, B. W. A., & Price, J. (2001). Implicit/explicit memory versus analytic/nonanalytic processing: Rethinking the mere exposure effect. *Memory and Cognition, 29,* 234-246.