Implicit Learning of Form-Meaning Connections

Janny Leung (janny@cantab.net)
Research Centre for English and Applied Linguistics,
University of Cambridge, English Faculty Bldg., 9 West Rd., Cambridge, CB39DP, UK

John N. Williams (jnw12@cam.ac.uk)
Research Centre for English and Applied Linguistics,
University of Cambridge, English Faculty Bldg., 9 West Rd., Cambridge, CB39DP, UK

Abstract

We report two experiments that used derived attention paradigms to examine implicit learning of associations between forms and meanings in semi-artificial languages. Both experiments examined learning of rules governing article usage, and the dependent measure was the time to locate objects referred to by article-plus-noun expressions in pictures. Learning was measured by slow-downs in search times when the learned associations were violated. Experiment 1 demonstrated implicit learning of associations between articles and the accompanying noun's thematic role in a sentence. Experiment 2 found no implicit learning of an association between articles and the position of the object referred to by the noun in a scene (only explicit learning effects were obtained). We explain this difference in terms of the possibility for interactions between implicit learning and grammatical knowledge.

Introduction

Studies of implicit learning have generally focused on the learning of rules and associations concerning forms in a single modality, for example letters in artificial grammar experiments, or spatial locations of visual stimuli in serial reaction time experiments. In the case of natural language systems learning such form-form associations might contribute to the acquisition of syntactic rules, but there remains the problem of learning how those forms relate to meanings. In this research we address the issue of whether people become sensitive to associations between forms and meanings without an intention to learn them, and without awareness of what they are.

The issue of learning form-meaning connections is usually considered in the context of vocabulary acquisition. Ellis (1994) argues that learning form-meaning connections is dependent upon declarative memory. For one thing, variables that are known to affect episodic memory also influence vocabulary learning, e.g. intelligence, depth of processing, meta-cognitive and mediational strategies. Gupta & Dell (1999) make a similar argument on purely computational grounds, pointing out that the arbitrariness of form-meaning mappings makes them very difficult to assimilate using the kinds of implicit (in their view, connectionist) learning mechanisms that underlie skill learning. Important evidence is also provided by research on amnesia. For example, amnesics such as H.M. (Gabrieli, Cohen, & Corkin, 1988) can learn novel word forms, as assessed by tests of implicit memory such as fragment completion and priming. But they seem to have great difficulty learning new vocabulary; for example they are unable to recognise or use words that they have been exposed to over long periods after the onset of amnesia. Amnesics are unable to recall specific experiences – they have impaired episodic memory. So their inability to acquire vocabulary implies that episodic memory is involved in learning form-meaning mappings, perhaps through conscious comparison processes between different occasions of use of a word (Ellis, 1994). However, more recent evidence suggests that amnesics can learn vocabulary after all (e.g., Vargha-Khadem et al., 1997). These studies count against the claim that vocabulary learning is necessarily dependent upon conscious memory for specific episodes of word use. Rather, the process of abstracting the meaning of words from experience may occur implicitly.

But even if we allow for implicit meaning abstraction in memory, there remains the question of how the requisite data gets into memory in the first place. For example, when a novel word is encountered, is it necessary to attend to the connection between that word and a (hypothesised) meaning? In the context of first language acquisition, Merriman (1999) argues that when the "principle of mutual exclusivity" is violated the child will "seek out, discover, and dwell on features that are uniquely associated with the second label" (*ibid*., p. 350). And in adult second language acquisition Doughty & Williams (1998) stress not only attention to form and meaning, but also their integration within the learning episode (*ibid*., p.245). These quotations imply that learners need to attend to relationships between forms and meanings. However, these arguments are merely speculative, and we know of no experimental investigations of implicit (as opposed to incidental) learning of form-meaning connections.

In researching these issues we have chosen to focus on function, rather than content, words. This makes it possible to examine potential interactions with prior grammatical knowledge and allows us to address the issue of whether some kinds of regularity are more readily learned implicitly than others because they capitalise on linguistic concepts that the learner already possesses.

The starting point was an earlier study (Williams, 2005) in which subjects first learned four determiner-like words:

gi, ro, ul, and ne. They were told that these functioned like the English definite article, except that they also encoded the distance between the speaker and the object, *gi* and *ro* being used for near objects, and *ul* and *ne* being used for far objects. They were then exposed to training sentences such as "I was terrified when I turned around and saw gi lion right behind me", "The researchers studied ul bees from a safe distance", "As I was passing I knocked over ro vase", "I asked the removal men to fetch ne boxes from the van". What the subjects were not told was that determiner usage also depended upon the animacy of the accompanying noun – *gi* and *ul* were used with animate objects, whilst *ro* and *ne* were used with inanimates. After training, subjects performed a generalisation test in which familiar words occurred in novel contexts and they had to choose between two possible determiners, e.g. "The art collector went to Greece to collect *ul / ne* vases", neither of which had occurred with that noun during training. Out of 44 subjects, 33 said that they were not aware of the relevance of animacy to article usage during training or testing phases, yet they were 61% correct in selecting the correct article in the test phase ($p < 0.001$). After a subsequent rule discovery task on the test items, 22 subjects were still unaware of the rule, yet their generalisation test performance was still above chance, 58% ($p < 0.05$).

The Williams (2005) experiment appears to provide evidence of implicit learning of one type of form-meaning connection. However, this may be a special case because the relevant meaning feature was part of the lexical representation of the nouns, and did not derive from the actual context of use of the word. To address the more general issue of learning form-meaning connections it is necessary to examine meaning features that derive from the context in which the word is used. This was one of the aims of the experiments reported here.

The experiments again used the words *gi, ro, ul* and *ne*, which were introduced to the subjects as article-like words. In Experiment 1 the target rule concerned the thematic role of the accompanying noun, whereas in Experiment 2 it concerned the spatial position of the noun's referent. Both experiments tested whether contextually, rather than lexically, derived meanings can become implicitly associated with novel words.

The methods we have adopted bear some similarity to the contextual cuing paradigm of Chun & Jiang (1998), and derived attention paradigm of Lambert and colleagues (e.g. Lambert & Sumich, 1996). Following work done within the "visual world" paradigm (Tanenhaus, et al, 1995) we assume a close linkage between language forms and attentional processes, but here we are using changes in attentional processes over time as indicators of learning. In addition, it may be through the effects of implicit learning on attention that the transition from implicit to explicit knowledge occurs (Jiminez, 2002). We therefore consider derived attention paradigms to be ideal for studying language learning processes.

Experiment 1

Method

Materials The learning target was a miniature article system. Participants were introduced to four artificial articles: *gi*, *ro*, *ul* and *ne*, where *gi* and *ro* were used before personal names referring to adults and *ul* and *ne* before personal names referring to children. What subjects were not told was that article use also depended on the thematic role of the noun phrase; *gi* and *ul* were used with agents and *ro* and *ne* with patients[1].

There were 98 training trials followed by 16 'control' trials and 16 'violation' trials. Each trial consisted of a distinct picture (like that in Figure 1) with an accompanying audio description, such as "Kiss ul Mary a boy on the face".

Figure 1: A sample picture used in the experiment

The subjects' primary task was to indicate whether the named individual (Mary in this case) appeared on the left or right hand side of the picture. To avoid some articles becoming more salient than others due to the position they appeared in, two sentence orders were used in the picture descriptions. Sentences followed this pattern:

V + NP1 + NP2 + PP

Where either NP1 or NP2 could be an agent or patient; NP1 was always a proper name and NP2 always a common noun; V was always in its infinitive form. Two possible sentence constructions to describe Fig. 1 were:

1. Kiss	ul Mary	a boy	on the face.
V	NP1(Agent)	NP2 (Patient)	PP
2. Kiss	ne David	a girl	on the face.
V	NP1(Patient)	NP2 (Agent)	PP

Pictures were selected so that an agent and a patient of the depicted action were clearly identifiable. Three common names for each gender were used with each possible article with equal frequency to avoid any input bias in article-name associations. The sound clips for the articles had been edited such that each of them lasted exactly the same amount of time, and the interval between the onset of the article and onset of the noun was 1 second. This provided a time window in which the article could be used to orient attention

[1]The terms 'agent' and 'patient' are used more loosely here than the specialist linguistic sense. For instance, 'agent' would include technical concepts such as agents, actors, causers, etc.

even before the noun was presented. Reaction time was measured from the onset of the article. We reasoned that if subjects had learned the relationship between the articles and thematic roles they would come to orient their attention to one or another individual in the picture even before the name was heard. For example, on hearing "Kiss ul .." in the context of Figure 1 they might anticipate that the following name refers to the kisser, i.e., the girl on the right of the picture in this case. If they heard "Kiss ne .." they would anticipate reference to the kissee, i.e., the boy on the left of the picture (note that all pictures were presented only once; we present this contrast here for expository purposes).

The critical data came from the last 32 trials of the experiment, divided into control and violation blocks (there was no division between these and the training trials from the subject's point of view). All of the pictures used in these blocks depicted either two adults or two children, so that only by knowing the mapping from articles to thematic roles would articles have any predictive value. In the control block the sentences respected the same system that was used in training, whereas in the violation block the mapping between articles and thematic roles was reversed (i.e. *gi* and *ul* were used with patients instead of agents). Items were rotated between the Control and Violation blocks over two groups of subjects so as to control for differences in item difficulty.

Procedure Twenty five subjects were tested, all of whom were native English-speaking university undergraduate or postgraduate students. None of them had knowledge of languages with highly developed case systems such as German and Latin. They first received vocabulary pre-training for the four articles, and they were told that the order of words in the sentences presented to them would not be as in English. Examples were given so that subjects became familiar with the articles and the sentence structure. They were given as much time as they needed for this pre-training.

For each picture, the subjects had to perform the following tasks (subjects had to click to proceed from one task component to the next):

Picture description: to describe the picture in their own words (this was to draw their attention to the action portrayed in the pictures);

Reaction time test: to click left or right on the mouse as quickly as possible as they listened to a picture description, once they could decide on which side of the picture the named individual appeared;

Sentence Reformulation: to put the sentence they heard in correct English order, retaining the article used (to stimulate concurrent active processing of the action portrayed and the article used).

A recall test was inserted between every two pictures, during which subjects had to repeat the sentence they reformulated in correct English for the last two pictures.

It is worth explaining that there were a few possible clues the subjects could rely on for the reaction time test. The most obvious clue was gender; where there was a man and a woman in the picture, subjects should have been able to react once they had heard the proper name. Having learnt about the adult-child distinction, where there was a child and an adult in the picture, the subjects could also react once they had heard the article, and before they heard the proper name. Adult-child pictures accounted for about 12 per cent of the trials. They served the purpose of keeping the subjects attending to the articles.

In a debriefing session subjects were probed for whether they had any feelings about when *gi* vs. *ro* and *ul* vs. *ne* were used, apart from the adult-child distinction. Subjects were asked to report any idea that came to mind. Based on their answer in this verbal report, they fell into 'aware' and 'unaware' groups.

The whole experiment took approximately one hour. A three minute break was inserted in the middle of the experiment to prevent fatigue.

Results and Discussion

Twenty out of twenty five participants remained unaware of the system at the end of the experiment. Subjects who were able to link the use of articles with concepts such as 'someone doing something to the other person' or 'one person takes an active role and the other is more passive' were all counted as aware, even if these were not the first guesses they made. Some subjects made up to four guesses. Many unaware subjects said it did not come to their mind at all that the two articles used for adults and children were used in different conditions, and were surprised when asked to comment on it. Those who thought there might be a differentiation most commonly reported that they speculated a gender distinction, but after some trials realised that it was not the case.

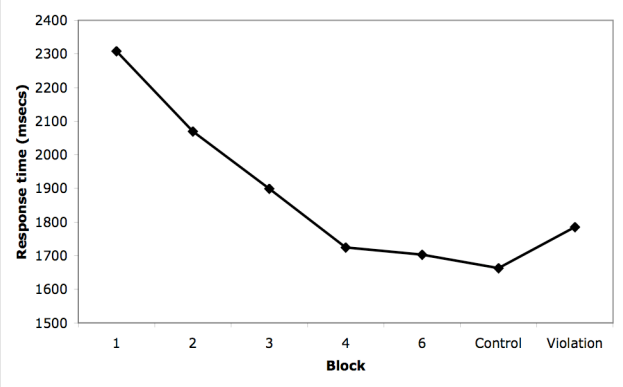

Figure 2. Reaction time over blocks in Experiment 1

Figure 2 shows the average reaction time against trial blocks for the unaware subjects. There is clearly a large task practice effect early in training, but what contribution does learning of the target system make to this learning curve? We can only gauge learning of the target system by looking at the reaction time difference between the violation block and the preceding control block. One can see in Fig. 2 that

there was a sudden increase in reaction time at the violation block, which went against the trend of the learning curve. The mean reaction time in the control block was 1663 ms, SE = 46 ms, and in the violation block it was 1785 ms, SE = 47 ms. The reaction times for these blocks were significantly different, $p<0.01$.

Error rates in the reaction time task were also informative. An error means that subjects clicked on the wrong side of the mouse (i.e. wrongly identified the named individual). If subjects had learnt the article system, error rate might increase in the violation block. The error rates in the control and violation blocks were 4.1% and 7.8% respectively, and the difference between them was significant, $p<0.05$.

Both a significant slow-down in reaction time and a higher error rate in the violation block indicated that the subjects were sensitive to the grammatical violation. In combination with their reported lack of awareness in the debriefing this suggests that they had implicitly learnt the agent-patient article system.

It is worth mentioning that the aware group showed a rather different pattern of results. Like that of the unaware group, the reaction time of the aware group showed a sharp decrease early in training, reflecting a task practice effect. Later in training, however, their average reaction time rose, before it gradually decreased again. The reaction time between the control and the violation blocks did not show a significant difference. Such a pattern could be explained by the fact that most 'aware' subjects had only become partially aware of the article system. Even though they had some idea that the agent/patient roles were relevant to article use, most of them could not yet match the articles with their respective roles, and thus they were still in the process of hypothesis formation and testing. This assumption is supported by the high error rate of the aware group when compared with the unaware group: the error rate of the aware group was about 5 times higher than the unaware group in the grammatical blocks, and was still 3 times higher during the violation block. Lambert & Roser (2001) report a similar lack of a violation effect in semi-aware participants.

Experiment 1 provides evidence for implicit learning of a connection between articles and thematic roles. The form-meaning connections were learned even though the relevant meanings were derived from the context, rather than the lexical representations of the nouns. This result is perhaps rather surprising in the light of the general assumption that learning form-meaning connections requires declarative memory, or at least the integration of forms and hypothesised meanings at encoding. It is also surprising in the context of skepticism about the possibility of acquiring even simple contingencies between stimuli without awareness. Lovibond & Shanks (2002) argue that, for example, there is no convincing evidence for classical eye blink conditioning in subjects who are not aware of the relationship between the unconditioned stimulus (puff of air) and conditioned stimulus (e.g. tone).

One reason for the appearance of an effect in this experiment may be that the target system relied on representations of thematic roles that are embedded in the language system, that were actively utilised in the sentence formulation task, and that were available for association with the articles. This is to say that there were interactions between implicit learning and grammatical knowledge. Likewise, in the Williams (2005) experiment the articles were associated with the animacy feature, which, cross-linguistically, also typically interacts with grammatical processes (see also Lambert & Sumich, 1996, for a demonstration of an animacy effect on learned derived attention). It therefore becomes important to see whether implicit learning effects can be obtained when forms are associated with concepts that are not typically grammaticised.

Experiment 2

Method

In this experiment, subjects first learned the pairing between the novel articles *gi* and *ul* and the English definite article 'the' (printed in red and green respectively), and *ro* and *ne* and the English indefinite article 'a' (in blue and grey respectively). Subjects learned these pairings, and then proceeded to the main task in which they saw pictures (e.g. Figure 3) and heard phrases (e.g. "ul crane").

Figure 3. An example picture from Experiment 2

The subjects' task was to indicate whether the object referred to by the phrase they heard was in the foreground or the background of the picture by pressing one of two response keys (arranged horizontally for half of the subjects, and vertically for the remaining half). Having made their response they repeated the phrase ("ul crane") and translated it into English ("the crane"). What they were not told was that objects in the foreground of the picture (and also below a fixation cross just below the centre of the display) were systematically referred to using *gi* (for 'the') and *ro* (for

'a'), whereas *ul* (for 'the') and *ne* (for 'a') were used with objects in the background (and also above fixation). Note that this meant that the novel words were also systematically associated with specific response keys (*gi* and *ro* with the foreground key, *ul* and *ne* with the background key). We reasoned that if the subjects learned this regularity then, on hearing *ul* for example they would be anticipating an object in the background of the picture, and be prepared to press the background response key. Should that regularity be reversed in a violation block, then a slow down in responses should be observed. The auditory phrases were again constructed so that there was a one second interval between onset of the article and onset of the noun.

There were 56 training stimuli, repeated over 3 blocks of trials. For the first two blocks the location of the target object was cued by a box which appeared coincident with the onset of the article in the auditory stimulus. There is evidence that task-irrelevant perceptual associations are only learned if they coincide with task-relevant stimuli (e.g., Seitz & Watanabe, 2005). Without the box as a cue the subjects would only orient their attention to the correct location after hearing the noun. However, in the final training block no box was presented and subjects had to find the object as quickly as they could for themselves (and hence could show sensitivity to the predictiveness of the article). The training trials were followed by 8 trials in which new pictures and phrases were used. These were followed by the final 24 critical trials, again using new pictures and phrases. For 12 of these, the control items, the novel articles were used in the same way as in the training, whilst for the 12 violation items a reversed mapping was used (e.g. *gi* and *ul* now used for background objects). The items were rotated between these two conditions across two presentation lists. In addition, half of the participants received the control items before the violation items, and half had the reverse order. In the debriefing, subjects were first asked if they noticed any regularities in the way that the articles were used. Specifically, given that there were two words for 'the' and two words 'a', why was sometimes one used and sometimes the other? Subjects who were unaware of the rule then worked through up to one block of training trials again, this time trying to work out the rule. They were told to think aloud as they did this.

Results and Discussion

A total of 16 subjects were tested, all of them native English-speaking university students. None of them reported awareness of the foreground/background rule during the training. However, when invited to search for a rule, 11 of them found the rule, in many cases after only a few examples. These subjects were adamant that they had not thought about the foreground/background distinction during the training phase, and did not notice any changes in the materials in the violation block. Indeed, they were rather surprised that they had not noticed such a simple regularity, especially given the correlation with their actual responses. The mean response time over control items was 2413 ms, SE = 89 ms (error rate 1.5%) and over violation items it was 2354 ms, SE = 79 ms (error rate 2.6%). The difference between the conditions was not significant, and in any case was in the opposite direction to that predicted. For the 11 subjects who were able to work out the rule in the rule search task the mean response times in the control and violation blocks were 2440 ms and 2367 ms respectively.

In order to check that the predicted violation effect could be obtained in aware subjects, another version of the experiment was constructed using the same items, but this time subjects were also required to recall the phrase used with each picture after two intervening items (items were presented in blocks of 3, followed by phrase recall cued by the picture + box). Out of 7 subjects tested, 5 became aware of the rule during training. Mean response times in the control and violation blocks were 2548 ms and 2845 ms respectively, $t(4) = 3.19$, $p < 0.05$. Thus, it cannot be argued that the present paradigm was insensitive to learning the correlation between the articles and the foreground/background distinction. Rather, learning effects were not obtained in the absence of awareness.

Conclusion

Clearly one must be cautious in making direct comparisons between Experiments 1 and 2 because of the procedural differences between them, forced in the main by the desire to keeps levels of learner awareness to a minimum. We note, however, that in many respects one would have expected implicit learning effects in Experiment 2 rather than Experiment 1. First, in Experiment 2 subjects were required to attend to the relevant meaning features in order to make their decisions, whereas in Experiment 1 those features remained implicit in the subjects' sentence formulation processes. Second, any violation effect could have occurred either through miscuing of attention, or through a change in the actual relationship between articles and response keys. Third, the majority of subjects in Experiment 2 rapidly became aware of the rule when invited to search for it, suggesting that it was an extremely simple regularity. So why then was it not learned implicitly?

Our hypothesis is that what is critical here is the potential relevance of the meaning features to grammatical systems. Features like animacy and semantic (thematic) role are closely related to grammatical processes, and might be expected to be relevant to the distribution of function words like articles. It is this notion of potential relevance which accounts for the greater availability of such features for associating with function words in these experiments. But a concept such as foreground / background (or more generally, near or far from the observer) whilst expressed in deictic terms such as "this" and "that" in English, for example, are less likely to interact with grammatical processes (for example, to control agreement within noun phrases). In the context of our experiments the situation becomes more like general vocabulary learning, which it has been argued is dependent upon explicit learning processes,

or at least attention to forms and hypothesised meanings at encoding. Alternatively, linguistic expectations may actually suppress learning in this situation because learners do not expect certain kinds of feature to be related to article systems. Prior linguistic knowledge might both facilitate and interfere with implicit learning.

Implicit / incidental learning research has only recently begun to re-consider natural language-like systems, but clearly as soon as one addresses this question one has to confront the issue of how prior linguistic knowledge influences the learning process. For example, Newport & Aslin (2004) have found that humans cannot learn non-adjacent dependencies between syllables under incidental learning conditions, whereas Tamarins can (Newport et al, 2004). One possible explanation of this is that the way in which phonological representations are structured actually prevents humans from being sensitive to non-adjacent dependencies at the syllable, but not the segment, level. In this case then, prior linguistic knowledge would be said to interfere with implicit learning. In general, findings such as these, and those reported here, begin to shed light on how domain-general statistical learning mechanisms interact with domain-specific knowledge.

Acknowledgments

This research was supported by a Sir Edward Youde Memorial Fellowship to Janny Leung.

References

Chun, M.M. & Jiang, Y. (1998). Contextual cuing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28-71.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 197-261). Cambridge: Cambridge University Press.

Ellis, N. (1994). Vocabulary acquisition: The explicit ins and outs of explicit cognitive mediation. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 211-282). London: Academic Press.

Gabrieli, J. D. E., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following medial temporal lobe resection. *Brain & Cognition, 7*, 157-177.

Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The Emergence of Language* (pp. 447-481). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Jiminez, L. (2002). Introduction: Attention to implicit learning. In L. Jiminez (Ed.), *Attention and Implicit Learning* (pp. 1-7). John Benjamins: Amsterdam.

Lambert, A.J. & Roser, M. (2001). Effects of bilateral color cues on visual orienting: revisiting William James' 'Derived Attention'. *New Zealand Journal of Psychology, 30*, 15-22.

Lambert, A.J. & Sumich, A.L. (1996). Spatial orienting controlled without awareness: A semantically based implicit learning effect. *Quarterly Journal of Experimental Psychology, 49A*, 490-518.

Lovibond, P.F. & Shanks, D.R. (2002). The role of awareness in Pavlovian conditioning: theoretical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behaviour processes, 28*, 3-26.

Merriman, W. E. (1999). Competition, attention, and young children's lexical processing. In B. MacWhinney (Ed.), *The Emergence of Language* (pp. 331-358). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Newport, E.L & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.

Newport, E.L., Hauser, M.D., Spaepen, G. & Aslin, R. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology, 49*, 85-117.

Seitz, A. & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences, 9*, 329-334.

Tanenhaus, M.K. et al (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, 268, 1632-4.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277*, 376-380.

Williams, J.N. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.