Attention, awareness, and noticing in language processing and learning

John N. Williams.
Department of Theoretical and Applied Linguistics,
University of Cambridge, U.K.

A slightly revised version of this manuscript appeared in J. M. Bergsleithner & S. N. Frota & J. K. Yoshioka (Eds.), *Noticing and Second Language Acquisition: Studies in Honor of Richard Schmidt* (pp. 39-57). Honolulu, Hawai'i: National Foreign Language Resource Centre. 2013.

Abstract

This chapter reviews current psychological and applied linguistic research that is relevant to Schmidt's landmark theoretical analysis of the concepts of attention, conscious awareness, and noticing. Recent evidence that attention and awareness are dissociable makes it possible to ask which of these constructs is required for either processing familiar stimuli, or learning novel associations. There is good evidence for processing familiar stimuli without conscious awareness, and even without attention. There is mounting evidence for learning without awareness at the level of understanding regularities, particularly when meaning is involved (semantic implicit learning). There is even some recent evidence for learning without awareness at the level of noticing form, for example learning associations between subliminal and liminal words. Thus, whilst attention does appear to be necessary for learning, awareness might not be. It is suggested that future research should probe the types of regularity that can be learned without awareness, and that this will shed more light on the nature of the underlying learning mechanism, and permit an evaluation of its relevance to SLA.

Author Bio

John Williams graduated in Psychology at the University of Durham, U.K., and then went on to do doctoral research at the Medical Research Council Cognition and Brain Sciences Unit, Cambridge, receiving a PhD from the University of Cambridge for work on semantic processing during spoken language comprehension. He spent two years as a post-doctoral researcher at the University of Padua, Italy, working on semantic access and word recognition. He worked as an English language assistant at the University of Florence for one year before taking up his present post in the Department of Theoretical and Applied Linguistics (formerly Research Centre for English and Applied Linguistics), University of Cambridge. His current main research interests are in the cognitive mechanisms of second language learning and second language lexical and syntactic processing.

Richard Schmidt's Noticing Hypothesis has been hugely influential in the field of SLA. Prompted by case studies of second language learners (Schmidt, 1983; Schmidt & Frota, 1986) Schmidt questioned what was then the dominant view of second language development as an unconscious, implicit, process. He noted that the presence of a form in the input did not guarantee acquisition, even after many years of exposure. A detailed analysis of his own output as a learner of Portuguese, the input he had received in the classroom, and the instances of language that he had noted in his diary showed that "R learned and used what he was taught if he subsequently heard it and if he noticed it" (Schmidt & Frota, 1986, p. 279). In other words, being taught a form was not sufficient, nor was exposure to it in the input, but he also had to notice the form in order for it to appear in his production. The term "notice" was used in its everyday meaning, " … a second language learner will begin to acquire the target like form if and only if it is present in comprehended input and "noticed" in the normal sense of the word, that is, consciously." (p. 311). In subsequent publications he backed up this intuitive idea with thorough reviews of the psychological literature on attention and learning (Schmidt, 1990, 1994, 2001) concluding that there is no evidence for learning forms without conscious attention. The attentional system filters the input and determines what becomes intake. However, he distinguished another sense in which awareness may be involved in learning – what he called "awareness at the level of understanding", which relates not to whether learners are aware of forms as such but whether they are aware of the rules that govern the distribution of those forms. He was open to the idea that this level of awareness may not be necessary for learning, although, again on the basis of the psychological literature of the time, remained sceptical about its contribution to adult second language acquisition.

Schmidt's ideas, his detailed analysis of the concept of attention, and his appeal to mainstream cognitive psychological research represents landmark work in the field of SLA. The Noticing Hypothesis fed into the development of pedagogic approaches that attempt to stimulate learning by encouraging an appropriate orientation of attention (Long & Robinson, 1988; VanPatten, 1996). He inspired researchers, including myself, to make direct connections between experimental work on cognitive learning processes and SLA. The Noticing Hypothesis constituted a clear, falsifiable, proposition that was open to empirical investigation using experimental methods.

In this chapter I will provide an update on the psychological background to the concepts underlying the Noticing Hypothesis. The theoretical concepts of attention and consciousness continue to be refined as more phenomena are explored experimentally, advancing our understanding of learning processes. Once we have characterised the learning processes then perhaps we will be in a position to evaluate their potential contribution to SLA.

What is attention? Dissociating attention and conscious awareness

William James (James, 1890) is often cited as expressing a common sense view of attention: "Everyone knows what attention is. It is the taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalisation, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others". Our intuition tells us that we are aware of what we pay attention to, "My experience is what I agree to attend to" (James, 1890, p. 403), and when we say we have not been paying attention to something we essentially mean that we have not been aware of it. We also intuitively identify the contents of awareness with "working memory" (Baars, 1988; Cowan, 1999), and consciousness as a "workspace" in which information from different sources can be integrated (Baars, 1988; Dehaene & Naccache, 2001). Schmidt (2001) concurs with this common sense identification of attention and awareness.

Attention without awareness

In fact recent research suggests that attention and awareness are separate, and potentially dissociable. Consider subliminal priming. If a stimulus is flashed very briefly and followed immediately by a masking pattern then, under the right conditions, an experimental participant will be unaware of what the stimulus was. Yet responses to a subsequent stimulus will be affected (Marcel, 1983; Perea, Dunabeitia, & Carreiras, 2008). This demonstrates processing without awareness. But does it show processing without attention? From the common sense perspective we might

think that it does – if the prime was not in awareness then it was not attended. However, there are now numerous demonstrations that top-down attentional factors affect processing of unconscious stimuli. Subliminal priming effects depend on attention being directed to the right moment in time when the prime is presented (Naccache, Blandin, & Dehaene, 2002), or to relevant kinds of information for supporting priming effects (Kunde, Kiesel, & Hoffmann, 2003; Spruyt, De Houwer, Everaert, & Hermans, 2012). A similar phenomenon has been found in studies of blindsight[1] (Kentridge, Heywood, & Weiskrantz, 2004). Despite claiming not to consciously see anything in their blind field, blindsight patients are able to make visual discriminations at above chance levels (e.g. indicate whether a bar in the blind field is horizontal or vertical). Just as for subliminal priming, this shows processing without awareness. However, it has also been shown that when provided with a valid spatial cue which they are aware of (e.g. an arrow pointing to a location in the blind field) their discrimination performance is enhanced in the blind field. So conscious attention can affect processing of an unconscious stimulus.

Consider how we orient attention. Clearly we are pre-programmed to attend to sudden and unexpected changes in the environment. The elephant crashing through the jungle will attract our attention, as will the police siren. In these cases we would say that we are aware of the stimulus, we orient our attention to it, engage in deeper processing, and decide what action to take. From the common sense view it would seem to be impossible to orient our attention to something that we are not even aware of. Yet this seems to be the case. In one study (Jiang, Costello, Fang, Huang, & He, 2006) arousing images were masked so as to make them completely invisible. Nevertheless, an image of a naked woman would be more likely to attract spatial attention than an image of a man, but only for heterosexual males (and the opposite for heterosexual females). This shows that quite fine distinctions can be computed for stimuli that are not conscious or focally attended, and that unconscious stimuli can direct conscious attention.

It seems from these examples that attention and consciousness should be thought of as different things (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Koch & Tsuchiya, 2007). It is therefore more precise to think of attention as a

[1] Blindsight patients are cortically blind, which is to say that the eyes and optic nerves are intact but there is damage to striate cortex (area V1). In some cases they can be blind on one side of the visual field and see normally on the other.

(limited) cognitive resource that when applied to a representation results in deeper and more elaborate processing. But just because something is attended does not mean that we are necessarily aware of it (as in the case of subliminal priming), and the attentional system can orient to stimuli that are not consciously perceived.

Awareness without focal attention

If awareness and attention can be dissociated, then it becomes possible to conceive of forms of awareness that depend minimally on attention. After all, focal attention is highly limited in capacity, and if we were only aware of the contents of focal attention then we would have a blinkered view of the world, a kind of tunnel vision. Yet our conscious experience seems to be much broader than this, as if it overflows the restrictions of focal attention. Driven by this intuition, Block (2007) has argued for a distinction between two forms of awareness: "access awareness" corresponds to the contents of focal attention, and is reportable. "Phenomenal awareness" extends beyond access awareness, is fleeting, and is not reportable. The concept of access awareness appears to correspond to Schmidt's notion of noticing, as being the reportable contents of focal attention. The intuition that there is more to our conscious experience than this perhaps gives rise to the idea that there might be forms of processing and learning that can occur outside of this narrow focus. But what is the evidence for phenomenal awareness?

In an experiment by Sperling (Sperling, 1960) people were briefly shown an array of 16 alphanumeric characters in a 4 by 4 grid. After it disappeared they could only report about 4 characters, yet they had the impression of seeing around 8. Thus, the contents of access awareness could be described as 4 recognised characters plus more items that were only processed to the level of being recognised as characters, but not identified. As one would expect from the common-sense view, only a limited number of items receive sufficient attention to be processed to the level of recognition, and there is minimal processing of the remaining items in the array. What is intriguing though is that if after the display disappears people are cued to report a specific line of the grid then they can do so with high accuracy (see also Landman, Spekreijse, & Lamme, 2003). This might be taken to imply that people are actually momentarily, and "phenomenally", aware of the identity of all of the characters in the display. But only a subset can receive sufficient attention to enter access awareness.

Not unsurprisingly, the claim that people are aware of (and identify) more than they can report is contentious, not least because it is difficult to prove (see the commentaries on Block, 2007). A simpler, and more traditional, explanation would be that the entire array is held in iconic memory, and that there is sufficient raw stimulus information to enable identification when cued. In this view, the contents of phenomenal awareness are fleeting and unreportable because they are unanalysed. But we can still retain the idea that consciousness overflows focal attention, consistent with our subjective experience.

Other experiments on visual perception have claimed that it is possible to achieve access awareness for objects outside of focal attention. For example, people can indicate whether complex pictures of natural scenes contain animals (Li, VanRullen, Koch, & Perona, 2002) or they can discriminate famous and non-famous faces (Reddy, Reddy, & Koch, 2006). However they cannot make what might appear to be simpler discriminations such as between large T's and L's, or red-green discs from green-red ones. Thus, there are clearly limitations on processing of non-attended stimuli, but at least in the case of vision, high level semantic descriptions of scenes ("gist") appears to be available to access consciousness. However, it is debatable whether such findings demonstrate awareness without attention. Cohen, Alvarez, & Nakayama (2011) suggest that in these experiments focal attention was not sufficiently taxed to deprive resources from the critical stimuli. They show that the ability to report natural scenes is drastically reduced when a highly demanding distracting task is employed. Interestingly, though, it is still not entirely eliminated, suggesting that a low level of access awareness for unattended stimuli is possible (or that no distracting task is completely distracting). See Cohen et al. (2012) for an argument that there is no double dissociation between attention and awareness.

The above studies attempt to explore awareness outside the focus of attention. The least radical interpretation of the findings is that there are iconic forms of memory that retain stimulus information in a relatively raw form, requiring minimal attention, but supporting conscious experience (phenomenal awareness). For certain classes of stimuli there can be access awareness outside of focal attention, but processing is not entirely attention-independent. After all it would be surprising if there were a complete firewall between information that is in focal attention and information that lies outside it.

All of this raises the possibility that there might be not only processing, but also learning, of stimuli outside of the attentional focus; that is, learning without noticing. Let's first consider processing.

Can there be processing without attention or awareness?

As mentioned above, Naccache, et al. (2002) showed that subliminal semantic priming depends upon attention. If the participant's attention is directed to a moment in time different from when the prime is flashed then no semantic priming effect is obtained. This would appear to show that there is no processing without attention (although as traditional subliminal priming experiments show, there can be processing without awareness). However, this is a case where the stimulus is presented extremely briefly and followed, or preceded, by a pattern mask, so there is limited information entering the system from the bottom up. But what about situations where the stimulus is not masked, and under normal conditions it would be easily perceptible, but we have no access awareness of it because our attention is occupied elsewhere? This is the kind of situation that might arise during language learning. For example, the relevant form is clearly present in the input but the learner does not notice it because their attention is focused elsewhere, perhaps on the meaning of the message.

The every day experience of mind-wandering provides a case in point. For example, whilst reading it sometimes happens that we start thinking about something that might (or might not) be totally irrelevant to the text. We become absorbed in our thoughts, yet our eyes continue to traverse the page. At some point we realise that our attention has wandered, a special state of consciousness that Schooler, et al. (2011) refer to as "meta-awareness" (the state of being aware that we are aware, or in this case, not aware). Our attention snaps back to the text, and perhaps we try to find the point where we lost track. Studies have shown that when mind-wandering during reading there are reduced event related potential[2] (ERP) effects, and eye movements lose their normal sensitivity to lexical variables such as frequency (Reichle, Reineberg, & Schooler, 2010). Clearly when attention is occupied elsewhere the degree of processing of the text is drastically reduced, and we have no access

[2] Event related potentials are brain responses to specific events and are measured using electroencephalography (EEG).

awareness of what the text was about, although we appear to have phenomenal awareness of the physical words on the page.

Similarly, in dichotic listening experiments people have to perform an attention-demanding task on the stimulus stream in one channel while ignoring the material in the other channel. Early research showed that participants were oblivious to even the most basic characteristics of the unattended channel, such as whether it is in a known language, or played forwards or backwards. Later studies showed that there may be some processing in the unattended channel, but it was difficult to rule out the possibility that these effects were caused by momentary switches of attention (Holender, 1986).

Although studies of mind-wandering and dichotic listening show that we have no memory for unattended material there still might be local processing effects that we are not aware of (and that might support some form of learning). Using dichotic listening Dupoux, Kouider, & Mehler (2003) looked for repetition priming effects from words in the unattended channel to lexical decisions on words in the attended channel. Under conditions when participants claimed not to have been aware of the words in the unattended channel there were no priming effects suggesting that there is no lexical access from unattended words. However, Dupoux et al embedded the prime words in time-compressed "babble noise" with no pause before or after the prime. Whilst segmentation of the words was possible under full attention, it is possible that the bottom-up stimulus information was insufficient for recognition without attention (akin to the situation under masking in the case of Naccache et al, 2002). In a similar paradigm, but using normal speech, Rivenez, Darwin, & Guillaume (2006) showed that there is no priming when the stimuli in the two channels have the same fundamental frequency (as was the case in Dupoux et al., 2002), but effects can be obtained when the frequencies are different. Most impressively of all, priming effects can be obtained when the participants have detected a target word in the attended channel that is coincident with the prime word in the unattended channel making it extremely unlikely that the prime was attended. Therefore, using natural speech that is perceptually distinguishable from the attended channel, but a task that makes attention-switching extremely unlikely, momentary lexical access from unattended words appears to be possible.

Studies of the attentional blink (AB) provide evidence for semantic processing of unattended words. The AB refers to the fact that after detecting one target in a very

rapid stream of stimuli there is a momentary lapse of attention, making it difficult to detect another target in the next 500 milliseconds or so, even though that target would ordinarily be perfectly visible (Luck, Vogel, & Shapiro, 1996). The question is, what kind of processing do "blinked" words receive? Numerous studies have provided evidence for semantic processing of words that are not detected during the AB (Frings, Bermeitinger, & Wentura, 2011; Luck et al., 1996; Martens, Wolters, & van Raamsdonk, 2002), and even for distracters that occur during the AB, which is surprising because they would be even less likely to be attended than targets (Harris & Little, 2010). Thus, there seems to be quite deep processing of stimuli that are not noticed, and which are not even attended, during the AB, even though they could cross the threshold into consciousness if attention were not occupied elsewhere. Dehaene, et al. (2006) refer to such stimuli as "pre-conscious" and suggest that because they produce relatively strong bottom-up activation they can attain deep levels of processing despite the absence of top-down attention.

To return to the language learning situation, just because the learner apparently has no access awareness, say, of a form, and their attention does not appear to have been directed towards it, this does not mean that they have not processed it to some level. Of course, the above studies employed familiar words in the participants' native language. It is a different matter whether there can be learning of new forms or novel associations between known forms without access awareness or focal attention.

Can there be learning without attention?

It is one thing to show short-term priming effects from unattended stimuli, it is another to show longer-term effects; that is, learning. The most basic form of learning is memory for a stimulus. Obviously, if an attentional manipulation has been successful participants will have little conscious, explicit, memory for unattended stimuli. And indeed research has consistently shown that conscious recognition memory is far superior for attended than unattended items, a phenomenon that has been referred to as "inattentional amnesia" (Rees, Russell, Frith, & Driver, 1999; Ruz, Worden, Tudela, & McCandliss, 2005). But as is the case in actual amnesia, might

implicit memory be revealed if indirect memory measures are used? Indeed, there is some evidence for this. Using a dichotic listening paradigm Bentin, Kutas, & Hillyard (1995) showed the usual substantial recognition memory advantage for attended over unattended words. However, of more interest were the responses to new words. Some of these were semantically related to 'old' words that had occurred in the dichotic listening task, and some were unrelated. The rate of false alarms (incorrectly judging as 'old') was higher for new words that were semantically related to words that had occurred in the dichotic listening task. Critically, this effect was as large when the new word was related to an unattended old word as an attended old word. They also found that facilitation of lexical decisions was as large for repetitions of unattended as unattended old words (repetition priming). These results suggest that the participants had equally strong implicit memory for unattended as attended words, despite having far superior explicit memory for the latter.

One study has even claimed that explicit memory for unattended words can be demonstrated provided that the memory test is sufficiently sensitive. In Hoffman, Bein, & Maril (2011) exposure to attended and unattended words was followed by a recognition memory task. As expected there was absolutely no discrimination between new words and unattended old words, apparently showing no conscious memory for unattended words. However, the participants were also required to rate their confidence as "high" or "low" for each decision. It was found that high confidence "no" responses were more likely for new words than unattended old words, and conversely, low confidence "no" responses were more likely for unattended old words than new words. Assuming that confidence ratings reflect conscious judgement knowledge (Dienes, 2008) then this study suggests that there can be conscious effects from memory traces left by unattended words, even if participants are unable to translate these into reliable recognition memory judgements.[3]

Both of the above examples only show "learning" in the sense of long term representational changes as a result of processing a known, but unattended, word. This form of learning is the basis for frequency effects, for the strengthening of

[3] In this and the previous case it is tempting to attribute the findings to attention-switching during the learning phase. However, this can not explain the results because in Hoffman et al. (2011) there was no evidence for actual recognition of unattended words, and in Bentin et al. (1995) performance on implicit tests was equally strong for unattended and attended words despite far superior recognition memory for the latter.

representations through exposure, and as such is an important aspect of learning a language (Ellis, 2002). But in language learning we are also, and primarily, concerned with learning new associations – either between existing phonological and orthographic representations to learn new word forms, between forms and meanings, and between forms and forms. Is there any evidence for this kind of learning without attention?

Experiments by Logan & Etherton (1994) show how attention affects learning of novel associations between known words. Participants were presented with pairs of words (one above the other, but both recognisable in a single fixation) and had to indicate if one of them belonged to a pre-specified semantic category. Some word pairs were periodically repeated, and the question was whether decision times would become faster for targets in repeated pairings than novel pairings. This was indeed the case demonstrating incidental learning of the word pairs. However, in another condition an arrow oriented attention to the target just before the word pair appeared meaning that participants never had to attend to the distracter word. Now the learning effect disappeared, even though the distracter word was presented close enough to fixation to be recognised. Presumably with no prior cue both words had to be attended in order to determine which one belonged to the pre-specified category. But with prior orientation of attention there was no learning of the association between attended words and unattended distracters.

Other experiments suggest that learning novel associations does not depend simply on whether a stimulus is attended, but on whether the appropriate dimension is attended. One has to pay attention in the right way for learning to occur. Jiménez & Méndez (1999) conducted a sequence learning experiment in which a stimulus moved around 4 positions on the screen and participants simply had to indicate its position by tapping corresponding keys. The stimulus appeared to be randomly one of * ? ! or x. In fact, its identity predicted the position of the next stimulus (e.g., a ! might predict that the next stimulus would occur in the right-most screen position). When participants only had to respond according to the position of the stimulus on the screen there was no learning of this regularity. Learning was only obtained when they had to also keep a running count of, for example, how many times either x or * appeared. Thus, learning the association between character identity and screen position only occurred when character identity was 'noticed', in Schmidt's sense. In fact, one can be even more specific because further analyses showed that what

participants actually learned was a relationship between being counted or not and the following possible screen positions - the specific associations between characters and screen positions were not learned. This is not to say that they did not also notice the identity of the stimulus, but the act of noticing whether the stimulus was counted was closer in time to the predicted event (the position of the stimulus on the next trial) than the act of noticing the identity of the character. This experiment provides a particularly good example of the relationship between noticing and learning, and how the timing of noticing events can be critical to learning.

Finally, in a more linguistically relevant context, Toro, Sinnett, & Soto-Faraco (2005) studied the human ability to segment 'words' from meaningless strings of syllables by tracking transition probabilities (e.g., the stream *tu-pi-ro-go-la-bu-bi-da-ku-pa-do-ti-go-la-bu-pa-do-ti-tu-pi-ro-bi-da-ku* is compose of the nonwords *tupiro, golabu, bidaku, padoti*). It had originally been claimed that this form of statistical learning does not require attention (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, Toro et al. showed that learning was drastically reduced under dual task conditions. For example, the pitch of some of the syllables was randomly increased by 20 Hz above the baseline frequency of 200 Hz and participants were required to press a button when they detected a pitch change. After 7 minutes of exposure performance on the word segmentation test was only 55%, not significantly different from the chance level of 50%, and significantly worse than the 64% achieved by participants who listened passively to the same stream. Even though the participants in the tone monitoring condition were attending closely to the input, they were attending to, and noticed, pitch and not syllables, and so they did not learn the transitional probabilities between them.

These experiments convincingly demonstrate the importance of attention to learning. In all cases there is a stimulus, or a stimulus dimension, that is clearly present, providing ample bottom-up stimulus information, but it is not receiving top-down attention. It is hard to believe that participants are not phenomenally aware of these unattended stimuli or stimulus dimensions. For example, in Toro et al. participants would presumably notice if the stimuli switched from syllables to simple tones of the same pitch. Information is present in phenomenal awareness, as a form of echoic memory, but it is not being attentionally selected for further processing and decision making because of the nature of the task that is set.

How do we square this with the evidence that there can be quite deep processing without attention, for example lexical access from unattended words in dichotic listening, or semantic processing during the attentional blink? If there can be deep processing and implicit memory without attention, why can there be no learning without attention? An obvious possibility is that activating existing representations in memory is one thing, whereas forming new connections in memory is another (Cowan, 1999). So in the Logan & Etherton (1994) experiment, under advance cue conditions a distracter word may well activate its lexical representation and meaning but for some reason this cannot form a lasting association with the focally attended word. Both need to be simultaneously attended for this to happen – or at least nearly simultaneously attended, as when in Jiménez & Méndez, 1999, people learn the association between 'counted' and the next screen position. Attention appears to provide a kind of "glue" for bonding stimuli together in memory.

Pacton & Perruchet (2008) provide a good example of this principle in the context of learning sequential dependencies. Previous research had shown that whereas people can track the dependencies between adjacent syllables (as in the example from Toro et al, above) they have trouble doing so when the dependent syllables are separated by a random syllable (Newport & Aslin, 2004). Pacton & Perruchet (2008) showed that associations between stimuli that are non-adjacent in the input stream can be learned if the learner's task requires them to attend to these stimuli at the same time (for example, in the case of digit sequences because they have to perform a calculation upon them). Unitization in the attentional focus does not have to be a response to specific task instructions. Gestalt principles of perceptual organisation can have a similar effect, as when the stimuli to be associated are similar and the intervening one is dissimilar, e.g., because they are consonants and the intervening item is a vowel (Newport & Aslin, 2004). As Pacton & Perruchet (2008) put it "associative learning is an automatic process that links together all the components that are present in the attentional focus at a given point" (*ibid*., p. 82). But whilst attention is clearly required for learning, what is the role of awareness?

Can there be learning without awareness?

Given that awareness and attention are dissociable, just because there can be no learning without attention does not mean that there can be no learning without awareness. Can people learn associations between stimuli without being aware of those associations? Or, as a more extreme scenario, can they learn novel associations between stimuli when they are not even aware of one of them (even though unconscious attention might be allocated to them)?

Schmidt (2001) made an extremely useful distinction between awareness at the level of noticing, which is having access awareness for forms in the input, and awareness at the level of understanding, which is having a conscious understanding of the relationships between forms. On the basis of a review of artificial grammar learning (AGL) research at the time Schmidt (1995) concluded that there was no evidence for learning without awareness at the level of understanding because such experiments actually provide no evidence for learning abstract rules, contrary to the original claims made by Reber (Reber, 1993). People might learn fragmentary chunks, and recurring sequences of letters at salient positions, but they do not learn the underlying grammar as such. Whilst Schmidt recognised that there may be forms of simple associative learning that can lead to learning without awareness at the level of understanding he was sceptical about the possibility of learning the kind of complex rule-based generalisations that underlie language.

Research in this area continues to support Schmidt's view. Even when AGL experiments appear to demonstrate acquisition of abstract knowledge through transfer (Altmann, Dienes, & Goode, 1995) this still does not demonstrate learning of the abstract grammar. What people appear to learn is the "repetition structure" of the training items; that is, salient patterns of alternation and doubling in the strings (Tunney & Altmann, 2001). Whilst this knowledge is abstract, it does not correspond to knowledge of the grammar as such. Nevertheless there is good evidence that whatever non-rule-based knowledge is being acquired in AG experiments is unconscious, as assessed through subjective measures (Dienes & Scott, 2005; Scott & Dienes, 2010). Serial reaction time experiments also provide good evidence of sensitivity to sequential regularities without awareness of what those regularities are (Jiménez & Méndez, 1999, is a good example). These all constitute examples of learning without awareness at the level of understanding but at a level below abstract generalisations.

In visual perception the contextual cuing paradigm reveals the ability to learn associations between targets and contexts without conscious understanding (Chun, 2000). People are required to locate a pre-specified target in a visual array containing what appear to be many randomly distributed distracters. For example, they might have to locate a rotated letter T in an array of randomly rotated L's. In fact some of the displays are repeated, and the remainder are random. As the task progresses, people become faster at locating targets in repeated as opposed to unrepeated displays, showing that they have learned the spatial context for these targets. Just as in Logan & Etherton's (1994) experiment with word pairs there is incidental learning of arbitrary associations provided that attention is at least momentarily allocated to the context (Jiang & Chun, 2002). The difference here is that learning was shown to be implicit because people had no conscious memory for the repeating displays – in a subsequent recognition test participants were unable to distinguish repeating displays from new displays. In another experiment people appeared to learn the association between the shapes of the distracters and the shape of the target, regardless of position (Chun, 2000). People implicitly learned which shapes went together, despite not being able to discriminate consistent from inconsistent pairings in a follow-up test. Thus, there is good evidence for learning associations between forms without awareness of what those associations are.

All of these experiments examine what are essentially form-level associations. Implicit learning at this level is presumably relevant to learning word forms, and multi-word units. From a statistical learning perspective, co-occurrences at the form level might correlate with, and bootstrap the acquisition of, higher-level grammatical distinctions (Saffran, 2003). But what about implicit learning beyond the form level and involving meaning? This is particularly important for functional and usage-based perspectives that emphasise the acquisition of form-function mappings (Ellis, 1998; MacWhinney, 1997; Tomasello, 2000).

Learning without awareness of meaning: Semantic implicit learning

Within experimental psychology there is growing evidence for meaning-based implicit learning, a phenomenon that one might refer to as 'semantic implicit learning'. In a variant of the sequence learning paradigm Goschke & Bolte (2007) showed that people can become sensitive to recurring semantic category sequences,

for example that a picture of an animal will follow a picture of a body part, without any awareness of the regularity. Variants of the contextual cuing paradigm have been developed in which target position is predicted by the semantic properties of the distracters – the oddness or evenness of digits (Goujon, Didierjean, & Marmeche, 2007), semantic category of words (Goujon, Didierjean, & Marmeche, 2009), or which kind of scene the target appeared in (bedroom, kitchen, etc., Goujon, 2011). These phenomena present a challenge to a strong form of the noticing hypothesis because generalisations are formed without awareness at the level of understanding, and at a level that goes beyond simple chunking of surface forms (see Paciorek & Williams, forthcoming, for further discussion of semantic implicit learning experiments).

There is also more direct evidence for semantic-based implicit learning effects within the domain of language. In Williams (2005) four novel determiners were injected into standard English. The participants were told that *gi* and *ro* were used with objects that were near to the speaker, and *ul* and *ne* with objects that were far. They then heard sentences such as "I was terrified when I turned around and saw gi lion right behind me", "The children threw sticks at ul monkey in the tree", "I knocked over ro cup and the coffee spilled on my book", "I couldn't read the title of ne book that was on the top shelf". What the participants were not told was that there was an additional regularity governing use of the novel determiners: *gi* and *ul* were used with living things and *ro* and *ne* with non-living things (see Rebuschat et al, this volume, for further procedural details). Participants who claimed to have been unaware of this regularity were nevertheless significantly above chance at choosing the correct determiner to use in novel article-noun combinations (generalisation items). The participants appeared to have developed some sensitivity to the correlation between the determiners and animacy without being aware of what that correlation was.

Of course, any claim for learning without awareness can only be as convincing as the means used to measure awareness. Post-experiment debriefing may be not the most sensitive measure as it is separated from the moment at which knowledge was used in actual test performance. Subsequent conceptual replications of the Williams (2005) study have addressed this issue. Chen et al. (2011) assessed awareness by requiring the participants to indicate whether each decision in the test phase was based on guess, intuition, memory, or rule. Across two experiments responses to

generalisation items that were based on guess and intuition combined were significantly above chance (and were 55% and 58% correct respectively). In contrast, in a conceptual replication that introduced many procedural changes, Hama & Leow (2010) asked participants to think aloud during both the training and test phases. There was no evidence of learning the animacy rule amongst unaware participants. However, results of another replication study by Rebuschat et al. (this volume) suggest that think aloud obliterates learning in this paradigm.

The above studies used judgement tasks in the test phase and these might have invited participants to draw on explicit knowledge. If veridical and unreported this will lead to spurious claims of learning without awareness; if erroneous it will reduce the observed learning effect. It is therefore better to use test tasks that reduce the potential contribution of conscious knowledge. To this end Leung & Williams (2012) adapted the Williams (2005) animacy system for a reaction time task. Essentially, on hearing a noun phrase, e.g. "ul monkey", participants had to indicate as quickly as possible by a button press whether the object being referred to is living or not. At the end of the experiment a block of violation trials were included in which the animacy rule was reversed. Participants who, in a post-experiment questionnaire claimed not to be aware of the correlation between determiners and animacy were still slower in the Violation block than the preceding Control block, suggesting sensitivity to animacy without awareness.[4]

Apart from animacy, other types of semantic distinction have been shown to support implicit learning. Using a reaction time methodology Leung & Williams (2011) showed implicit learning of a correlation between novel determiners and the thematic role of the following noun, this time in a task that did not require an explicit decision based on the hidden regularity. Using a method based on Williams (2005) Guo et al. (2011) showed implicit learning of semantic prosody, the tendency for some words to take collocates with a positive or negative connotation (e.g. 'cause' tends to take negative collocates, whereas 'promote' tends to take positive ones).

[4] Interestingly, the participants who reported awareness of the rule showed a distinctive reaction time profile in that they were able to make their animacy decisions before the noun had been heard (there was a one second interval between determiner onset and noun onset). This pattern was not observed in any of the "unaware" participants, suggesting that what people reported in the post-experiment questionnaire reflected their use of conscious knowledge in the task.

The methodology used in all of these semantic implicit learning studies draws participants' attention to the relevant forms. In most of them their attention is not explicitly drawn to the relevant aspects of meaning, although in some cases it is (the reasons for this variation are presently unclear). What is critical, though, is that the participants do not appear to be aware of the correlation between forms and certain aspects of meaning, and yet this affects their performance in detectable ways. Thus there is learning at the level of understanding, but without awareness. This contrasts with the general outcome of research within the AGL tradition where it has been very difficult to show that participants implicitly learn much by way of the underlying abstract structure of the grammar. Maybe though there is something more "natural" about forming generalisations over form-meaning correlations than over, for example, sequences of letters generated by a finite state grammar. The issue of naturalness will be revisited in the Conclusion.

Learning associations without awareness at the level of noticing

Finally, we return to experimental psychology research to consider what appears to be a case of learning without awareness at the level of form, or in Schmidt's terms, without awareness at the level of noticing. Participants in Alonso, Fuentes, & Hommel (2006) saw target words on a screen and had to indicate whether they were animals or items of furniture. Each word was preceded by a subliminal, masked, prime word which was one of two category labels – all animals were preceded by the word BODY, and all items of furniture by the word PLANT e.g., BODY - dog, PLANT - chair. The question was whether participants would learn the association between the categories (body part – animal, plant – furniture). This was tested in a subsequent lexical decision task, again using subliminal primes, in which the order of the categories was reversed. The prime ANIMAL was followed by exemplars of body parts, e.g. ANIMAL – hand, and FURNITURE was followed by exemplars of plants, e.g., FURNITURE – rose. Facilitation was obtained relative to the opposite pairings (e.g. ANIMAL – rose, FURNITURE – hand). So people had learned an arbitrary bidirectional association between the categories (BODY – ANIMAL, PLANT – FURNITURE) that generalised to words not encountered in training. This occurred without awareness at the level of noticing one of the words during either the learning or testing phases. This effect has subsequently been replicated (Custers & Aarts,

2011). Using a different methodology Schlaghecken, Blagrove, & Maylor (2007) also showed that the predictiveness of masked cues can be learned without awareness of what those cues are

On the face of it, these results appear to challenge the idea that attention is required for learning novel associations. Crucially, though, note that the claim here is not that learning occurred without attention, merely that it occurred without awareness at the level of noticing form. Novel associations can be learned even when one of the relevant forms is not consciously perceived, but an appropriate attentional orientation is still necessary (Custers & Aarts, 2011).

Finally, these studies revealed evidence of an interesting qualitative difference between learning with and without awareness. In another of Alonso et al.'s (2006) experiments no masking was used and all words were clearly visible. Although all of the participants would have had awareness at the level of noticing the primes, it was found that about half of them were aware of the regular relationship between primes and targets (e.g. that BODY was always followed by an animal). Interestingly, it was this aware group, who had achieved awareness at the level of understanding, who did *not* show a priming effect when the order of the categories was reversed in the lexical decision test phase. This effect was only obtained for the unaware group. Aware participants appear to learn unidirectional associations, whereas unaware participants learn bidirectional associations. Custers et al. (2011) showed that even participants who were unaware because primes were masked would only learn unidirectional associations after their attention had been "tuned" to process pre-target events as predictive. They argue that it is not awareness, but attention, that plays a causal role in learning.

Why were bidirectional associations only learned without awareness? Alonso et al. (2006) suggest that aware participants learn that BODY *predicts* animals. But this does not mean that ANIMAL will predict body parts, hence there is no priming when the direction of the association is reversed in the test phase. On the other hand, unaware participants merely learn that BODY and ANIMAL words "go together", and so presentation of either one will activate the other and produce priming. One could regard the first kind of learning as cognitive and inferential, and the second as associative. Learning with and without awareness might lead to qualitatively different kinds of knowledge (Jiménez, Vaquero, & Lupianez, 2006).

Conclusion

The common sense view of attention, expressed for example by James (1890), accords with our everyday experience – we are aware of what we attend to. What we attend to is processed deeply and is learned, and what we do not attend to is not. However, psychological research suggests that things are not so simple. Attention and awareness can be dissociated – attention affects processing of stimuli that we are not aware of, and there may be forms of awareness that depend minimally on attention (phenomenal awareness). Attention does not determine all that is processed – there can be unconscious recognition, and even semantic processing, of stimuli that are not even attended, let alone conscious. There is even some evidence for implicit memory for non-attended words, demonstrating a form of learning without attention that could support the reinforcement of representations with repetition. However, in the SLA context we must bear in mind that many of these phenomena concern words in the native language. Perhaps then they tell us more about the automaticity of processing than the non-selectivity of attention, and we may wonder at what level of fluency such effects could be detected in the L2. But in principle, it does appear that non-attended stimuli can lead to lasting changes in representations.

When it comes to learning new associations, rather than reinforcing existing representations, then there is ample evidence that attention to the relevant forms, or more specifically the relevant dimensions of forms, is necessary. But is awareness, or what Schmidt referred to as "awareness at the level of noticing", also necessary? Whilst the assumption within psychology and SLA appears to be that it is, some recent evidence suggests that new associations might be formed even when one of the stimuli to be associated is not consciously recognised. Although this may challenge the idea that awareness is necessary for learning, it does not diminish the role of attention (as dissociable from awareness). Note also that the effects again relate to native language words for which lexical access can be assumed to be highly automatic.

What about awareness of associations, as opposed to the stimuli themselves, what Schmidt referred to as "awareness at the level of understanding"? At the level of learning associations between forms, as chunks, or sequences, the psychological

literature on implicit learning provides numerous examples of learning simple regularities without awareness of what they are. Implicit learning of this type underpins statistical learning more generally. However, Schmidt's scepticism was directed more at learning generalizable rules. In the area of ordering rules underlying artificial or semi-artificial languages evidence for learning without awareness remains elusive. However, there is mounting evidence from a variety of paradigms that people can show sensitivity to meaning-based generalisations without awareness of what those generalisations are, although attention to the relevant stimuli during learning is clearly necessary.

A central observation in Schmidt's diary study was that "R subjectively felt as he was going through the learning process that conscious awareness of what was present in the input was causal" (Schmidt & Frota, 1986, p. 281). The experimental evidence reviewed here suggests that learning is still possible without conscious awareness, at least of regularities, and possibly of forms themselves. Of course, processing is deeper and more sustained with awareness, and learning effects are much larger. The resultant knowledge can be intentionally controlled, and used flexibly. In contrast, even though effects of learning without awareness are detectable, the effects tend to be small. They produce unconscious biases in performance; for example, slightly faster times to locate a target in a display (Chun, 2000), a slight bias towards accepting items that respect a semantic regularity (Chen et al., 2011) or slightly slower semantic decision times when a regularity is violated (Leung & Williams, 2012). It is no surprise then that in comparisons of broadly "implicit" and explicit teaching methods that do not employ sensitive performance measures, the explicit methods appear to be more effective (Norris & Ortega, 2000; Spada & Tomita, 2010). At the present time, though, the main motivation for examining learning without awareness is theoretical, and through that to understand the contribution that it could make to learning problems such as SLA that are of a scale that cannot be investigated in the lab.

Of course the question whether there can be learning without awareness continues to be debated within both psychology and applied linguistics (Hama & Leow, 2010). Some psychologists continue to deny the possibility, essentially on the grounds that learning consists, minimally, in the association of ideas (Shanks, 2007), and in the extreme, even simple associations are not mere mechanical connections, but have propositional content and attitudes (Mitchell, De Houwer, & Lovibond,

2009). In this view, even something as simple as classical conditioning requires awareness of the relevant predictive relationships (Lovibond & Shanks, 2002). Note, however, that such research examines completely novel and arbitrary associations, such as pairings of a tone and puff of air, which do not make contact with prior knowledge structures. In contrast, most of the examples of learning without awareness described above involve meaningful material, stimuli that automatically activate representations in a rich knowledge base. This is particularly true of the cases of semantic implicit learning. Thus, it is important to investigate the relationship between learning and awareness in environments that are more naturalistic than those typically studied in psychology experiments.

Once we accept the relevance of existing knowledge structures then we need to explore how they constrain the learning process. For example, there is evidence that not all semantic distinctions are equally learnable. Leung & Williams (2012) found no implicit learning of a correlation between determiners and the relative size of two objects (e.g. "gi" would occur with the smaller object, regardless of its absolute size). And using the judgement task based on Williams (2005), Chen et al. (2011) found no learning of a correlation between determiners and the size of the object relative to a dog. Both of these systems are linguistically unnatural suggesting that, within language at least, semantic implicit learning is constrained. The constraints may derive from prior knowledge of universally encodable semantic distinctions (Bickerton, 2001) or from general, and still potentially universal, cognitive factors such as concept availability. Thus, we move away from the question of whether implicit learning occurs at all to a more theoretically interesting one of whether there are innate or more general cognitive predispositions for learning certain patterns within language. Clearly, if we are interested in why some linguistic correlations are spontaneously absorbed by the mind and others not then we are going to have to employ paradigms that tap into implicit, as opposed to explicit, learning processes. When we know more about the areas of language that appear amenable to implicit learning, then we will know where to look for its effects in real world language learning situations.

References

Alonso, D., Fuentes, L. J., & Hommel, B. (2006). Unconscious symmetrical inferences: A role of consciousness in event integration. *Consciousness and Cognition, 15*, 386-396.

Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*, 899-912.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Bentin, S., Kutas, M., & Hillyard, S. A. (1995). Semantic processing and memory for attended and unattended words in dichotic-listening: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 54-67.

Bickerton, D. (2001). Okay for content words, but what about functional items? Commentary on Bloom: How children learn the meanings of words. *Behavioral and Brain Sciences, 24*, 1104-1105.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*, 481-548.

Chen, W. W., Guo, X. Y., Tang, J. H., Zhu, L., Yang, Z. L., & Dienes, Z. (2011). Unconscious structural knowledge of form-meaning connections. *Consciousness and Cognition, 20*, 1751-1760.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*, 170-178.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science, 22*, 1165-1172.

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences, 16*, 411-417.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge: Cambridge University Press.

Custers, R., & Aarts, H. (2011). Learning of predictive relations between events depends on attention, not on awareness. *Consciousness and Cognition, 20*, 368-378.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*, 204-211.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*, 1-37.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research, 168*, 49 - 64.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: distinguishing structural knowledge and judgment knowledge. *Psychological Research, 69*, 338-351.

Dupoux, E., Kouider, S., & Mehler, J. (2003). Lexical access without attention? Explorations using dichotic priming. *Journal of Experimental Psychology: Human Perception & Performance, 29*, 172-184.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning, 48*, 631-664.
Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.
Frings, C., Bermeitinger, C., & Wentura, D. (2011). Inhibition from blinked category labels: Combining the attentional blink and the semantic priming paradigm. *Journal of Cognitive Psychology, 23*, 514-521.
Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 394-406.
Goujon, A. (2011). Categorical implicit learning in real-world scenes: Evidence from contextual cueing. *Quarterly Journal of Experimental Psychology, 64*, 920-941.
Goujon, A., Didierjean, A., & Marmeche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition, 15*, 257-275.
Goujon, A., Didierjean, A., & Marmeche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 50-71.
Guo, X., Zheng, L., Zhu, L., Yang, Z., Chen, C., Zhang, L., Ma, W., & Dienes, Z. (2011). Acquisition of conscious and unconscious knowledge of semantic prosody. *Consciousness and Cognition, 20*, 417-425.
Hama, M., & Leow, R. P. (2010). Learning without awareness revisited. *Studies in Second Language Acquisition, 32*, 465-491.
Harris, I. M., & Little, M. J. J. (2010). Priming the semantic neighbourhood during the attentional blink. *Plos One, 5*.
Hoffman, Y., Bein, O., & Maril, A. (2011). Explicit memory for unattended words. *Psychological Science, 22*, 1490-1493.
Holender, D. (1986). Semantic activation without conscious identification in dichotic-listening, parafoveal vision, and visual masking - a survey and appraisal. *Behavioral and Brain Sciences, 9*, 1-23.
James, W. (1890). *The Principles of Psychology*. New York: Holt.
Jiang, Y., & Chun, M. M. (2002). Contextual cueing: reciprocal influences between attention and implicit learning. In L. Jiménez (Ed.), *Attention and Implicit Learning* (pp. 277-296). Amsterdam: Bejamins.
Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences, 103*, 17048-17052.
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., Vaquero, J. M. K., & Lupianez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 475-490.
Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia, 42*, 831-835.
Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences, 11*, 16-22.

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*, 223-242.
Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43*, 149-164.
Leung, J. H. C., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually derived meanings. *Studies in Second Language Acquisition, 33*, 33-55.
Leung, J. H. C., & Williams, J. N. (2012). Constraints on implicit Learning of grammatical form-meaning connections. *Language Learning, 62*, 634-662.
Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, 99*, 9596-9601.
Logan, G. D., & Etherton, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1022-1050.
Long, M. H., & Robinson, P. (1988). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 3-26.
Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature, 383*, 616-618.
MacWhinney, B. (1997). Second language acquisition and the Competition Model. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 113-142). Mahwah, New Jersey: Lawrence Erlbaum Associates.
Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197-237.
Martens, S., Wolters, G., & van Raamsdonk, M. (2002). Blinks of the mind: Memory effects of attentional processes. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 1275-1287.
Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183-246.
Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science, 13*, 416-424.
Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.
Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417-528.
Paciorek, A. W., & Williams, J. N. (forthcoming). Semantic implicit learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*.
Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 80-96.
Perea, M., Dunabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58*, 916-930.
Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge*. Oxford: Oxford University Press.

Reddy, L., Reddy, L., & Koch, C. (2006). Face identification in the near-absence of focal attention. *Vision Research, 46*, 2336–2343.
Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science, 286*, 2504-2507.
Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*, 1300-1310.
Rivenez, M., Darwin, C. J., & Guillaume, A. (2006). Processing unattended speech. *Journal of the Acoustical Society of America, 119*, 4027-4040.
Ruz, M., Worden, M. S., Tudela, P., & McCandliss, B. D. (2005). Inattentional amnesia to words in a high attentional load task. *Journal of Cognitive Neuroscience, 17*, 768-776.
Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science, 12*, 110-114.
Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*, 101-105.
Schlaghecken, F., Blagrove, E., & Maylor, E. A. (2007). Incidental learning of S-R contingencies in the masked prime task. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1177-1188.
Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and Language Acquisition* (pp. 137-174): Rowley, MA: Newbury House.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.
Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 165-209). London: Academic Press.
Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-63). Honolulu: Hawai'i: University of Hawai'i Press.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge: Cambridge University Press.
Schmidt, R., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326): Rowley, MA: Newbury House.
Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences, 15*, 319-326.
Scott, R. B., & Dienes, Z. (2010). Knowledge applied to new domains: The unconscious succeeds where the conscious fails. *Consciousness and Cognition, 19*, 391-398.
Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60*, 291-309.
Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*, 263-308.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*, 1-29.

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition, 122*, 91-95.
Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*, 156-163.
Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*, B25-B34.
Tunney, R. J., & Altmann, G. T. M. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning Memory and Cognition, 27*, 614-639.
VanPatten, B. (1996). *Input Processing and Grammar Instruction: Theory and Research*. Norwood, New Jersey: Ablex Publishing Corporation.
Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.