Attention and Awareness: Evidence from Cognitive and Second Language Acquisition Research

Bandar Al-Hejin[1]
Teachers College, Columbia University

ABSTRACT

Over the past two decades, researchers in the field of second language acquisition (SLA) have become increasingly interested in the concept of attention. The role of attention has significant implications for theories of second language input, processing, development, variation, and instruction. Most of the literature on attention also addresses the concept of awareness. The two concepts are inherently connected but can be operationally distinguished. This review focuses on attention and awareness with respect to their definitions, theories, and the empirical evidence of their role in learning. Emphasis will be on research from cognitive psychology. This is followed by a brief review of SLA research relevant to Schmidt's Noticing Hypothesis (1990). The discussion suggests that: (a) Schmidt's definition of attention, as detection among other things, disregards the possibility of learning without attention; (b) attention facilitates SLA; (c) awareness is not necessary for converting input to intake; and (d) awareness can also facilitate SLA.

INTRODUCTION

Over the past two decades, researchers in the field of second language acquisition (SLA) have become increasingly interested in concepts traditionally associated with cognitive psychology such as memory, learnability, and connectionism. Ellis (2002) points out, "We are now at a stage at which there are important connections between SLA theory and the neuroscience of learning and memory" (p. 299). The concept of *attention* has become especially important because of its crucial role in so many aspects of SLA theory such as input, processing, development, variation, and instruction. Most of the literature on attention also addresses the concept of *awareness*. The two concepts are inherently connected but can be operationally distinguished. The following review will focus on attention and awareness with respect to their definitions, theories, and the empirical evidence of their role in learning. Emphasis will be on research from cognitive psychology. This is followed by a brief review of SLA research relevant to Schmidt's Noticing Hypothesis (1990).

[1] Bandar Al-Hejin holds a Master of Education in Applied Linguistics from Teachers College, Columbia University. He is currently an EFL instructor at the Institute of Public Administration in Riyadh, Saudi Arabia. His research interests include the effects of modality, especially digital media, on what language learners notice in the input. He is also interested in spatial (deictic) perception in Arabic. Correspondence should be sent to Bandar Al-Hejin, P.O. Box 205, Riyadh 11141, Saudi Arabia. E-mail: hejinb@ipa.edu.sa.

DEFINITIONS AND THEORY

Attention and related terms such as *consciousness*, *noticing*, *awareness*, and *understanding* are sometimes used interchangeably in the literature, making it difficult to compare theories and results from different studies (Schmidt, 1994a). This may be due to the inherent subjectivity in defining these concepts. In fact, in classical psychology, attention and consciousness are often viewed as two sides of the same coin. As Carr and Curran (1994) point out, "if you are conscious of something, then you are attending to it… and if you are attending to something, then you are conscious of it" (p. 219). Moreover, everyday use of the term *conscious* has a variety of overlapping meanings such as *awake*, *aware*, and *deliberate*. The reason for this overlap, as the following discussion will illustrate, is that these concepts are inherently connected, and one concept often entails the other. To clarify the distinctions between these terms, and for the purpose of consistency, this review will adopt definitions by Schmidt (1994a), which seem to be the most widely cited in SLA literature.

Schmidt (1994a) identifies four dimensions to the concept of consciousness. The first is *intention*, which refers to a deliberateness on the part of the learner to attend to the stimulus. Intention is often associated with intentional versus incidental learning. Chomsky (1975), for example, argues that children's acquisition of their first language is always incidental since children never really choose to learn their mother tongue. The second dimension of consciousness is *attention*, which basically refers to the detection of a stimulus. The third dimension is *awareness*, which refers to the learner's knowledge or subjective experience that he/she is detecting a stimulus. Awareness is often associated with explicit versus implicit learning, since learners may or may not be aware that they have acquired a new structure (e.g., children generally seem unaware of the complex syntactic rules they acquire). The fourth dimension of consciousness is *control*, which refers to the extent to which the language learner's output is controlled, requiring considerable mental processing effort, or spontaneous, requiring little mental processing effort.

Turning to attention, Tomlin and Villa (1994) suggest there are four conceptions of attention in SLA. One is that of attention as a limited capacity system. The idea being that the brain may be presented (through the sensory system) with an overwhelming number of stimuli at any given time, and it seems impossible to process them all. The limitations of attention refer not only to the amount (or duration) of attention that may be given to a single stimulus but also to the number of stimuli that may be attended to simultaneously. This leads to a second conception of attention, namely that it constitutes a process of selection. The overwhelming amounts of incoming stimuli force the attentional system to be selective. The third conception of attention, already touched on under consciousness, is that it involves controlled rather than automatic processing of information. The underlying assumption here is that some tasks require more processing effort, and hence a higher degree of attention, than others. A person may therefore perform two tasks at the same time, especially if one requires automatic processing (low attention). By the same token, it is more difficult to perform two tasks if both require controlled processing (high attention). The fact that controlled processing of two simultaneous tasks *is* sometimes possible led researchers to develop a fourth conception of attention, which is that it must involve a process of coordination among competing stimuli and responses. In this process, attention must be established, maintained, discontinued, and redirected in order to perform different actions.

Posner and Petersen (1990) describe attention in terms of three networks: *alertness*, *orientation*, and *detection*. Alertness refers to a general state of readiness to receive input. The higher the level of alertness, the faster the speed of selecting information for processing will be. If selection is too quick, however, the quality of processing may suffer. Orientation refers to the alignment of attentional resources to a particular stimulus from among a host of stimuli. Orienting attention to a stimulus facilitates the processing of that stimulus. Posner, Walker, Friedrich, and Rafal (1987) propose that orientation is made up of three mechanisms: disengaging from a stimulus, shifting to a new one, and re-engaging with a new stimulus. Orientation differs from alertness in that a learner might for example be ready to learn (alertness) but not know whether to focus on form or meaning (orientation). Detection is probably the most important network in attention; it refers to the cognitive registration of a stimulus. Once a stimulus is detected, it becomes available for further processing. Although detection does not necessarily imply awareness, Schmidt (2001) suggests using the term registration to refer to stimuli that are detected without awareness.

Awareness, as indicated, refers to an individual's subjective experience of a stimulus or cognitive content. Allport (1988) suggests that three conditions must be met in order for a person to be aware of a given experience. First, the person must show a behavioral or cognitive change as a result of the experience. For example, a learner might begin using *–ed* endings as a result of having been exposed to input that targets the past tense. Second, the person must report that he/she was aware of the experience at the time it took place. For example, the learner might report having been aware of *–ed* endings in the verbs at the time of exposure. Finally, the person must be able to describe the experience. For example, the learner must be able to articulate the morphological rule underlying the regular past tense. Leow (2000) adopts a less strict definition of awareness that requires only the first two conditions to be met. He calls this low awareness. High awareness is achieved when all three conditions are met.

To understand the role of awareness in learning, it is necessary to distinguish between learning and knowledge. Reber (1989) points out that these concepts have been confused even by experts. He defines implicit learning as "the process by which knowledge about the rule-governed complexities of the stimulus environment is acquired independently of conscious attempts to do so" (p. 219). This contrasts with explicit learning, where the learner is aware of, and actively involved in, processing the input. Learning (a process) is then distinguished from knowledge (a product). Paradis (1994) identifies two types of knowledge. One is implicit knowledge, which is acquired without awareness, unavailable to conscious memory (even after acquisition), and put to use spontaneously without conscious control (e.g., linguistic competence). The second type is explicit knowledge, which is knowledge that the learner is aware of and can access on demand (e.g., metalinguistic knowledge of grammar). Some argue that no interaction between the two forms of knowledge is possible (Krashen, 1982; Paradis, 1994; Truscott, 1998). Others argue that such interaction is possible at the level of learning, which implies a crucial role for noticing (Carr & Curran, 1994; Ellis, 1994; Schmidt 2001; Robinson, 1995). In any case, it is important to distinguish these concepts, because assuming for example that awareness is necessary for learning should not necessarily suggest that such explicit learning is going to result in explicit knowledge (even though the two are commonly associated with each other).

So how does noticing fit into all of this? According to Schmidt (1994b) noticing refers to the "registration [detection] of the occurrence of a stimulus event in conscious awareness and subsequent storage in long term memory…" (p. 179). In terms of the dimensions discussed

earlier, we might represent Schmidt's definition as follows: *noticing = detection + awareness*. However, since it is impossible to be aware of something without detecting it, we might as well simplify the equation to *noticing = awareness* (according to Schmidt). Schmidt is careful to distinguish noticing from *understanding*, which he defines as "recognition of a general principle, rule or pattern" (1995, p. 29). Understanding represents a deeper level of awareness than noticing which is limited to "elements of the surface structure of utterances in the input" rather than underlying rules (Schmidt, 2001, p. 5). The above definitions of consciousness, attention, and awareness are summarized in Figure 1 (for an alternative illustration showing the entailments of different attentional concepts see the Appendix).

FIGURE 1
The Four Dimensions of Consciousness and the Factors affecting them Based on Schmidt (1994a) and Allport (1988)

Consciousness
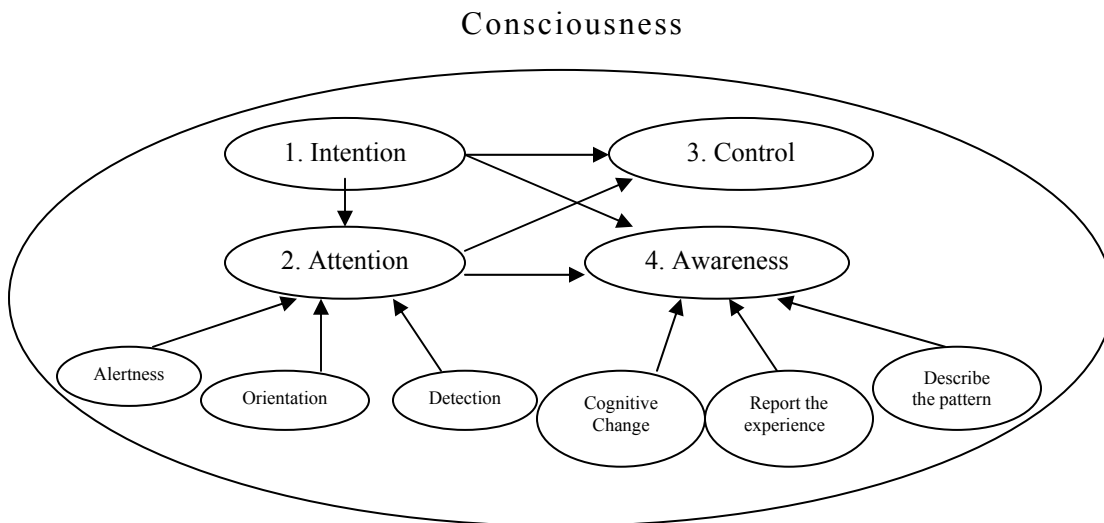
Robinson (1995) discusses two categories of attentional theories in cognitive psychology, both of which incorporate memory, an aspect he believes has been ignored in SLA literature. The first category, *filter theories*, addresses the notion of selection in attention (i.e., which of the incoming stimuli will be selected for further processing). The assumption in all filter theories is that there is only one channel of attention. Accordingly, simultaneous performance of two tasks can only be achieved through rapid alteration of attention between them. One such theory is the "bottleneck" model (Broadbent, 1958, as cited in Robinson, 1995) where input is: (a) registered by the senses, (b) selected (filtered), (c) detected, and (d) entered into short-term memory. This early selection model assumes that only form can be noticed at the selection stage, and that meaning is analyzed once the input has passed the filter. Other early selection models propose that both form and meaning are analyzed before detection, but this would involve substantial mental processing of the input before noticing even takes place. Finally, there are filter models that propose that selection does not take place until the input has entered short-term memory (late selection). In this view, selected input is rehearsed and unselected input is forgotten.

Rehearsal can involve *data-driven processing*, which refers to maintaining instances of the input in memory, or *conceptually-driven processing*, which refers to the activation of schemata from long term memory (Jacoby, 1983). A summary of filter models is illustrated in Figure 2.

FIGURE 2
Early and Late Selection/Filter Models of Attention Adapted from Best (1992) cited in Robinson (1995)
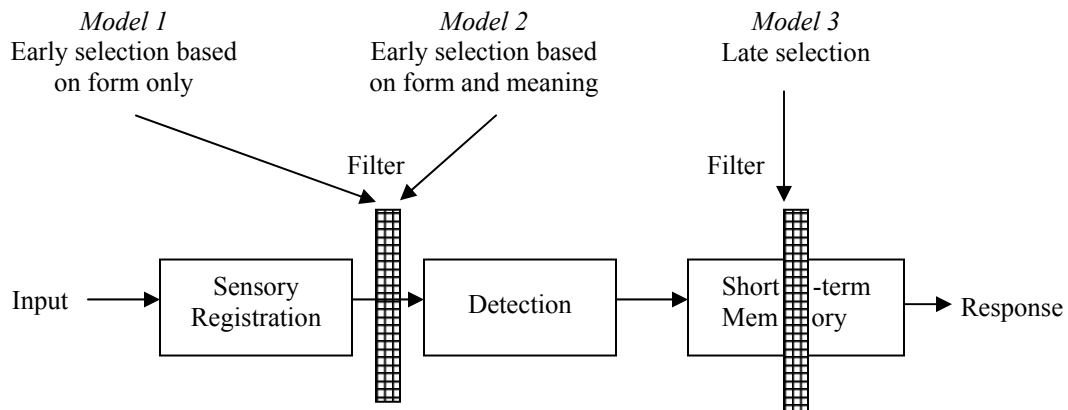
The second category of attention theories are *capacity theories*, which differ from filter theories in that they do not assume that the stimuli are necessarily competing for attentional resources. Instead of describing attention as a filter, the metaphor here is that of a spotlight, which can be narrowed (focal attention) and broadened (global attention). Key factors in capacity theories are the voluntariness of attention and the attentional demands of the task itself. Wickens (1980) proposes that attentional resources are best viewed as multiple pools that lie along three intersecting dimensions: (a) cognitive activities versus response processes, (b) processing codes of spatial versus linguistic activities, and (c) processing modalities (auditory versus visual perception and vocal versus manual responses). He argues that attentional demands become more difficult when simultaneous tasks are drawing on the same attentional pools such as, for example, trying to hold two conversations at the same time. In such cases serial processing is necessary as the focus of attention shifts between the two tasks. In other cases, where the attentional demands differ, parallel processing is possible, but the quality of performance may suffer as a result. For example, most people can drive while talking on their cell phones, but many countries outlaw that practice because it seems to distract focal attention from driving. Wickens points out that individuals may differ with respect to memory and parallel processing ability, suggesting that such variables need to be accounted for in attentional studies.

Most theories assume that attention plays a positive role in processing input because it involves detection and selection. However, attention may also be viewed as an inhibiting mechanism that prevents items from being processed. According to Schmidt (2001), "research on inhibitory processes is probably the most active and theoretically interesting work within attention theory at the present time" (p. 22). Absence of the inhibitory mechanisms of attention is

best illustrated in attention deficit disorders, which prevent individuals from focusing their attention on a given stimulus. In such cases, individuals are unable to resist or inhibit potentially attention-capturing interference, suggesting that inhibition is an important part of the attentional process.

EVIDENCE FROM COGNITIVE SCIENCE

The Nature of Attention

Before discussing studies that address the role of attention in learning, I will briefly review some research that informs about the nature of attention itself. This will provide a better understanding of some of the attentional theories and definitions mentioned in the previous section, specifically, attention as a limited capacity system, late selection models, the three networks of attention (alertness, orientation, and detection), and attention as an inhibitory mechanism.

The notion that attention is a limited capacity system was demonstrated by Sperling (1960). Five participants were presented with nine letters arranged in three rows for 50 milliseconds (ms) per trial. When asked to recall all letters, participants' average recollection was 4-5 letters. In order to rule out the possibility that incomplete recollection was due to visual limitations, participants were directed to recall a particular (upper, medium, or lower) row *after* the letters were removed from the display. This procedure led to almost perfect recollection of each row. These findings suggest that attention is limited in terms of orientation rather than detection. That is, despite having detected (seen) all nine letters, attention could only be oriented to a limited part of the input (one row). This also suggests that selection takes place after detection in accordance with late selection models (see Figure 2).

Neuroanatomical evidence seems to support the description of attention as three distinct but interrelated networks (Posner & Petersen, 1990). Monitoring electrical activity and blood flow in the brain using event related potentials and positron emission tomography has enabled researchers to relate attentional networks to certain regions of the brain. Accordingly, alertness, orientation, and detection were found to be associated with increased activity in separate but anatomically connected areas of the brain (Posner & Petersen, 1990; Posner & Rothbart, 1992).

The notion of attention as an inhibitory mechanism has been demonstrated in *negative priming* studies. In such studies, a participant is simultaneously presented with two stimuli. The participant is asked to respond (attend) to one stimulus while ignoring the other for a number of consecutive trials. If that participant is then asked to respond to the previously ignored stimulus on a subsequent trial, his/her response time is often slower than it was to the previously attended stimulus, which is not surprising because the participant may have developed practice effects with the original stimulus. What is surprising is that the response time of the same participant to a completely new stimulus is often faster than it is to the previously ignored stimulus. This phenomenon, a delayed reaction time to previously ignored stimuli, is referred to as negative priming and it seems to reflect an inhibitory mechanism of attention that *tags* stimuli as irrelevant for subsequent processing, making it more difficult (time consuming) to attend to such stimuli when asked to do so.

In a series of experiments, DeSchepper and Treisman (1996) showed that negative priming can occur with novel shapes as stimuli. In one experiment, participants (n=10) were presented with a pair of overlapping shapes that were colored red and green next to another shape that was colored white. They were asked to ignore the red shape and decide as quickly as possible if the green shape matched the white one. Response was given by pushing one of two buttons labeled "same" and "different". Shapes were presented such that the ignored red shape (negative prime) would reappear on a subsequent trial as a target shape in green. Results showed that response time was slower to negatively primed shapes, suggesting that an inhibitive effect was present. In another experiment with different groups (n=20, 24, 30, 32), DeSchepper and Treisman found that negative priming effects lasted up to a month after exposure. However, not all participants showed evidence of such effects.

The above studies are some of the few that provide insight into the nature of attention. They also provide evidence to support some of the aforementioned theories that attempt to define attention as being a limited resource as well as a faculty that can function below the level of consciousness (i.e., without awareness).

Attention and Learning

Early experiments investigating the role of attention in learning tended to be selective attention studies, which require participants to attend to one stimulus while ignoring another. These studies typically use dichotic listening tasks, in which participants wear headphones that present them with a different aural stimulus in each ear. Tasks are designed to heighten attention to one ear while depleting it in the other. After exposure, an assessment is made as to how much of the unattended stimulus was detected. For example, Moray (1959) asked 12 participants to shadow[2] prose passages that were presented in one ear while listening to 35 repetitions of seven words in the other ear. Results showed that the suppressed stimulus was not noticed as participants were unable to recall any of the seven words. Another dichotic listening study with two participants by Norman (1969) also found that recollection of the unattended words was very poor. However, some recollection was possible when the time between exposure and testing was very short, suggesting that attention may not be necessary for cognitive registration in short-term memory. Using a slightly different experimental design, Allport, Antonis, and Reynolds (1972) found that performance on divided attention tasks was dependent on the nature of both tasks. In one experiment 6 participants were asked to shadow a passage under one of three conditions: (1) listening to words in the other ear; (2) viewing words; (3) viewing pictures. Visual stimuli were displayed on a projector screen. Results showed that recognition of the unattended stimulus was poor in all three settings. However, there were statistically significant differences between each condition. Condition (1) was most affected by divided attention, followed by condition (2), and condition (3). This suggests that divided attention was most difficult when the two tasks shared the same modality such as the aural tasks in condition (1). In another experiment, five music students were asked to sight-read and perform piano music while shadowing an aural passage as a distraction. This time, the distracting stimulus seemed to have little or no effect on the performance of the primary task. These findings seem to confirm Wickens' (1980) model of

---

[2] *Shadowing* refers to repeating the prose out loud while listening

attention as multiple pools, which proposes that divided attention is best coordinated under contrasting task demands and modalities.

One drawback of dichotic listening studies is that they cannot rule out the possibility that participants paid some attention to the suppressed stimulus. Even when instructed to ignore one stimulus and shadow the other, participants may momentarily switch attention (Schmidt, 1995). Another drawback is that these studies often rely on introspective measures (recall and recognition), which may not be reliable considering that a participant may attend to something but fail to remember it on the posttest. This is illustrated in a more innovative dichotic listening study by Eich (1984). Sixteen participants shadowed a passage that was presented in one ear and word pairs were simultaneously presented in the other. The second word in each pair was a homophone (e.g., *fair* and *fare*); the first word was strongly associated with the less common interpretation of the homophone (e.g., *Taxi- Fare*). Following exposure, a recognition test (an introspective measure) was used to assess retention of the suppressed stimulus. As in previous studies, recognition of the suppressed stimulus was poor. However, a subsequent spelling test on homophones (a performance measure) revealed that participants were more likely to produce the less common interpretation with homophones to which they had been exposed than homophones to which they had not. In other words, the performance measure seems to have tapped into attended but forgotten (or unnoticed) knowledge, which the introspective measure failed to catch. Accordingly, performance measures (also called objective measures) have been more common in recent cognitive research.

Recent research has also relied more on divided attention studies, in which participants are asked to perform two tasks simultaneously. One of the tasks usually requires concentrated attention, making it extremely difficult to attend to the other task. The point of this measure is to ensure that attention is not oriented to the suppressed stimulus. The attended stimulus is often a Serial Reaction Time (SRT) task, an experimental design pioneered by Nissen and Bullemer (1987). The four experiments in their study are widely cited as evidence that attention is necessary for learning. Moreover, a number of subsequent studies have adopted their methodology such as Cohen, Ivry, and Keele (1990) and Curran and Keele (1993) (see below).

Nissen and Bullemer (1987) presented participants with a light that appeared in one of four possible locations on a computer monitor. The task was to press a keyboard button corresponding to each light as soon as it appeared. Reaction time was recorded by a computer. Each participant was presented with eight blocks of 100 trials. Each trial consisted of 10 light sequences appearing either randomly (random condition) or in a certain order (repeating condition). Participants were unaware of the sequencing, which made them perceive each block as a continuous series of 100 trials. There were 12 participants in all groups. In the first experiment, the repeating group and the random group conducted only the SRT task. Participants in the random group were chosen such that their mean SRT was equivalent to those in the repeating group. Results showed that the random group improved their SRT by 32 ms, which was attributed to practice. The repeating group showed a more significant improvement of 164 ms, which suggests that they had learned the sequence and were able to predict where the next light would appear. The second experiment included three new groups. Two groups were presented with an SRT task with a repeating sequence as before. But for one group, the computer generated either a high or low tone after each light, and participants were asked to report the number of low tones they had heard at the end of each block. The third group was presented with the same dual task except that the sequence was random. Another difference is that all three groups were tested on predicting where the next light would appear after the fourth block, which

would provide evidence of more explicit learning. Measures of both response time and prediction did not show evidence of learning in the repeating dual-task group, whose performance was inferior to the repeating single-task group. Moreover the difference in response time between the repeating and random dual-task groups was not significant. In other words, dividing attention seemed to have a significant negative effect on performance, suggesting that attention was necessary for learning. The third experiment compared the performance of the repeating dual-task group on a subsequent repeating single-task with the performance of a group that had no previous experience. Interestingly, the more experienced group did not outperform the novices, confirming that they had not learned anything from their previous experience under divided attention conditions. (Nissen and Bullemer's fourth experiment will be discussed in the next section on awareness.)

Different results were found in a similar SRT study by Cohen et al. (1990). Their methodology was a variant of Nissen and Bullemer's (1987) dual versus single task and random versus repeating sequence paradigm. One of the questions the researchers wanted to address was whether the nature of the sequence affected learning in a dual task. This was achieved by presenting experimental groups with three types of sequences (n= 12 in all groups). One sequence contained five items, none of which were repeated in a sequence. That is, each item had a unique association with the following item within a sequence. By assigning numbers to represent the position of the lights from left to right we can represent this type of sequence as follows: 1 5 2 4 3. The second type of structure contained six items, each one of which occurred twice (e.g., 1 3 2 3 1 2). In this structure no unique association existed between two adjacent items, making it impossible to use pairwise linear association to learn the structure. The third type of structure was a hybrid version of the first two, containing two unique associations and two ambiguous associations. In contrast to the original study, Cohen et al. found that structured sequence learning was possible in dual tasks as long as the structure type was linear (types 1 and 3). Hierarchical structures (type 2), however, could not be learned when attention was divided. This may have significant implications for language learning since most current theories of syntax subscribe to a hierarchical structure of constituents (Radford, 1997). Another question addressed in this study was whether increasing distracter difficulty affects learning of structured sequences. This was achieved by varying the number of low tones to be counted in different experimental groups. Results showed that increasing the attentional demands of the distracting stimulus did not have a significant effect on performance, suggesting that learning linear structured sequences may require only residual attention. However, Cohen et al. admit that different results may have been obtained if the dual tasks were sharing the same modality (i.e., tapping into the same attentional pools).

Another variant of Nissen and Bullemer's (1987) SRT study was conducted by Curran and Keele (1993). Although the study's focus was on transfer effects between declarative and procedural knowledge, their results have frequently been interpreted as evidence for learning without attention. In one experiment, one group was provided with an explanation of the sequence that the lights would follow (intentional group) and the other group was not (incidental group) (n=15 in each group). Both groups underwent single and dual task trials to assess whether what had been learned under single-task conditions would transfer to dual-task conditions. As might be expected, the intentional group outperformed the incidental group on the single task, but the two groups showed the same limited amount of learning under dual task conditions. This seems to confirm Cohen et al.'s findings that non-attentional learning is possible. It also suggests that the type of knowledge acquired under non-attentional conditions (procedural knowledge)

differs from the type of knowledge acquired when attention is paid (declarative knowledge), since participants appeared unable to utilize their declarative knowledge in non-attentional contexts.

Based on an experimental paradigm developed by Reber (1967), Whittlesea and Dorken (1993) assessed learning of miniature artificial grammars (MAGs)[3] under three attentional conditions. In condition one, participants (n=26) were misled to believe that the purpose of the experiment was to test their short-term memory of three-digit numbers. Between the presentations of numbers, participants were asked to repeat letter strings out loud as a distraction task. In condition two, the letter strings were the only task, and participants had to determine whether a particular letter had been repeated in other parts of the string. In condition three participants were asked to memorize the letter strings. Following all conditions, participants had to provide grammaticality judgments of strings using the same letters or different letters. Results showed above-chance judgments for all groups when the original letters were used. However, when different letters were used, and knowledge of the MAG was necessary to give the correct response, results showed no evidence of learning under condition one (dual task). This would suggest that attention is necessary for learning even linear structures. The contradiction between these findings and those of Cohen et al. (1990) and Curran and Keele (1993) may be due to differences in the modality of the input (lights versus letter strings).

Modality differences between the attended and the unattended tasks seem to be a general problem with divided attention studies. If attention is perceived using the aforementioned analogy of a spotlight that can be focused and broadened, then we must consider the possibility that the *unattended* stimulus may have been partially attended, especially when the two stimuli are of different modalities and draw on different attentional pools. An attempt was made to address this issue in a divided attention study by Kellogg (1980). The study used mental multiplication as the attended stimulus and face recognition as the unattended stimulus. Participants (n=156) were asked to: (a) look at the center of a slide projector screen showing a face with each slide, (b) visualize two numbers being read to them with each slide, (c) mentally multiply the two numbers, and (d) write down their answer. Participants were told that the more they visualized the multiplication problems, the more accurate their calculations would be. Kellogg hoped that such visualizations would control for the difference in modalities problem by forcing visual attention to the numbers and away from the faces. A surprise recognition test was administered five minutes after the exposure. Participants showed better than chance recognition of the faces presented, suggesting that attention was not necessary for long-term memory.

The studies reviewed so far seem to provide conflicting evidence as to the necessity of attention in learning. This may be due to the methodological difficulty of demonstrating a complete absence of attention. According to Schmidt (1995), studies claiming to show evidence of learning without attention are only showing that less attention leads to less learning. He argues that the participants in Kellogg's study could easily have ignored the instructions and momentarily attended to the faces presented. Since there was no way of verifying that participants followed instructions, it seems quite possible that some attention may have been paid to the suppressed stimulus. Schmidt also suggests an explanation for the conflicting evidence in the SRT studies. He points out that the light sequences used in Nissen and Bullemer's (1987) study had longer patterns (10 positions) and were of a more complex nature since some of the

---

[3] MAGs consist of strings of letters (e.g. ELFENAD, OLFAPID) generated by an underlying system of linear rules or *grammar*.

elements could be followed by more than one subsequent element (e.g., position 2 could be followed by either positions 4 or 3). In Curran and Keele (1993), the light sequences followed shorter patterns (6 positions) and each element could only be followed by one subsequent element (e.g., 2 can only be followed by 4; 1 can only be followed by 3). Assuming more complex stimuli require more attention, it would make sense that participants were unable to learn Nissen and Bullemer's complex patterns with divided attention but were able to do so with Curran and Keele's simpler patterns. This seems to be confirmed by Cohen et al.'s (1990) findings that only simpler linear structures could be learned without attention.

But does this mean that some learning without attention is possible? Even Curran and Keele (1993) admit that "When we refer to one form of leaning as nonattentional, we do not wish to imply that no attention whatsoever is used on the primary task. Undoubtedly, subjects must in some sense attend to a visual stimulus to make a response" (p. 190). The concept of nonattentional learning is both methodologically and theoretically problematical. Methodologically, it seems difficult if not impossible to demonstrate a complete absence of attention when learning takes place (Schmidt, 1995). Very often it seems to be awareness rather than attention that is absent in such cases. Since learning necessarily requires detection, and since detection is the essence of attention, learning without attention is also theoretically impossible (Truscott, 1998).

Awareness and Learning

As mentioned, awareness refers to the learner's knowledge or subjective experience that he/she is detecting a stimulus; implicit learning is learning that takes place without awareness. Before addressing the issue of implicit learning, we will briefly touch on the more basic issue of implicit perception. Is it possible to perceive something without being aware of it? This question is addressed in a widely cited study by Marcel (1983), which used semantic priming as a measure for perception. Semantic priming refers to a person's tendency to respond quicker to a target word (e.g., *student*) when it is preceded or primed by a semantically related word (e.g., *teacher*) than when it is preceded by a semantically unrelated word (e.g., *yogurt*). Marcel exploited this tendency using primes that were presented (masked) below an objectively defined threshold of awareness. Individual thresholds were previously established by testing different degrees of masking and asking participants if they were aware of the presence of a word (masking refers to obscuring the target word with different letters). Participants (n=12) were primed by asking them to focus on the masked words for a certain amount of time. Next they were asked to read aloud as quickly as possible words presented to them on a screen. Priming effects were measured by response time to the target word. Results showed evidence of semantic priming even when the prime word was masked below the threshold of awareness. This led Marcel to conclude that perception without awareness was possible. Note that these findings cannot be interpreted as perception without attention since participants were clearly attending to the screen.

Using a variant of the SRT methodology, Hartman, Knopman, and Nissen (1989) found further evidence of learning without awareness. Instead of responding to a pattern of lights, participants were presented with a series of words on a computer monitor. In the first experiment, participants were instructed to repeat each word aloud as quickly as possible when it appeared. Reaction time was measured from the appearance of the word on the screen to the

onset of a verbal response from the participant. The experimental group (n=15) was presented with a 10-word repeating sequence of four words as follows: MUSIC, RULER, LADY, OCEAN, LADY, RULER, MUSIC, LADY, RULER, OCEAN. The control group (n=15) was presented with the same four words in a random fashion. Neither group was informed about any order in the words. After the task, participants were asked if they had noticed a pattern, and if so, to describe it. Participants unable to identify more than three consecutive words from the list were classified as "unaware". In the second experiment, participants were presented with a different set of words: MAPLE, HAMMER, SALMON, ROBIN, SALMON, HAMMER, MAPLE, SALMON, HAMMER, ROBIN. Instead of repeating each word, participants were instructed to say the semantic category to which each word belonged (i.e., ROBIN: BIRD, HAMMER: TOOL, SALMON: FISH, and MAPLE: TREE). The results in both experiments suggest that better performance (faster reaction time) was associated with awareness of the repeating word pattern. However, even unaware participants showed significant learning when asked to respond with the semantic category in the second experiment. Schmidt (1995) takes issue with the researchers' classification of unaware learners in this study. He claims that it is "hard to understand" why Hartman et al. chose to classify participants who could not report more than three consecutive words as being unaware (p. 22). However, Hartman et al.'s measure of awareness was not haphazard, as Schmidt seems to suggest, but based on empirical evidence from Willingham, Nissen, and Bullemer (1989, as cited in Hartman et al., 1989), which shows that objective performance measures of awareness (e.g., guessing what the next word will be) of participants reporting less than four items in a 10-item sequence consistently matched those reporting no awareness at all. Hartman et al. acknowledge that participants classified as unaware "may have had some sense that a pattern was present, [but] they were unable to make use of that knowledge consciously" (p. 1072). Nevertheless, Schmidt's objection to the term *unaware* is well taken because a participant's ability to report three consecutive words from the list may in fact constitute some awareness (rather than none). Accordingly, future research might consider using classifications such as *less aware* rather than *unaware*.

Most of the abovementioned attentional studies also tried to assess the effects of awareness as being distinct from attention. Following SRT trials, Nissen and Bullemer (1987) questioned participants as to whether they had noticed a sequence, and if so, to report when they had noticed it. Nine out of 12 participants in the single-task repeating group reported an awareness of the sequence, and 11 out of 12 in the dual-task repeating group did not report any such awareness. Since the performance of the single-task group was superior to that of the dual-task group, Schmidt (1995) cites these awareness reports as evidence that "there was a *very* strong relationship between awareness and learning in these experiments" [his emphasis] (p. 21). However, this conclusion overlooks an important detail. As Nissen and Bullemer point out, "The response times of most normal subjects showed evidence of some learning well before they reported that they noticed the sequence" (p. 29). This suggests that learning was in fact possible without awareness because learning seems to have preceded awareness.

Similar results were found by Cohen et al. (1990). Following the primary SRT task, participants performed another task requiring them to predict where the next light would appear by pressing certain buttons (Nissen & Bullemer, 1987). Unlike the SRT task, which requires implicit knowledge, the prediction task requires participants to use explicit knowledge. By comparing performance at the end of the SRT task with performance at the beginning phase of the prediction task, it was possible to infer how much explicit knowledge had already been available and thereby assess awareness. This method has two advantages over questionnaires and

verbal reports. One is that it does not lead the participant into awareness. The second advantage is that it avoids potential guessing and forgetting. Results showed that some participants who had mastered the sequence implicitly were able to predict the positions of the next light, but others were unable to do so despite showing evidence of implicit learning. This led the researchers to conclude that awareness may not be necessary for memory and learning.

Nissen and Bullemer's (1987) fourth experiment is believed to provide even stronger evidence of the dissociation of awareness and learning. The same SRT procedures were conducted with 6 participants suffering from anterograde amnesia resulting from Korsakoff's syndrome (a condition that severely inhibits access to explicit or conscious memory). They found that participants were able to learn the structured sequences. Since Korsakoff patients have no access to declarative memory, their performance is often explained as learning without awareness. Carr and Curran (1994) have gone as far as claiming that structural learning by amnesics is "the gold standard of implicit learning" (p. 216). Schmidt (1995), however, claims that learning by anterograde amnesics is actually unreliable evidence of implicit learning. He points out that normal persons are able to report their level of awareness some time after the exposure, but amnesics cannot provide such reports due to their condition, which makes it impossible to assess their awareness in the first place. That is, according to Schmidt, the inability to recall the state of awareness at the time of exposure does not necessarily entail that the person was unaware at that time. However, even if the amnesic was aware at the time of exposure, we might consider that his/her awareness differs in one form or another from the type of awareness that enters long-term memory in normal persons. A normal person may be aware, forget, and then remember again because his/her awareness is linked to explicit long-term memory with lasting effects that may be realized at a later time. The amnesic's awareness, on the other hand, is fleeting and may only have immediate effects that cannot be remembered or revisited at a later time. It must be linked to implicit memory if it is to have any long-term effect at all. In any case, the question of whether forgotten awareness can affect learning is one that needs further investigation.

Awareness was also assessed in Curran and Keele's (1993) SRT study. One group of participants (n=14) was explicitly informed about the pattern of the lights before the trial. Another group (n=30) was administered an awareness questionnaire that asked them if they thought the lights appeared at random locations or according to some pattern. Those who thought there was a pattern were asked to describe it. Participants were then classified as *more aware* if they were able to correctly identify four of the six sequence positions, and *less aware* if they could identify three or fewer. Results showed that participants who were explicitly informed did better than the *more aware* group (n=19), who in turn did better than the *less aware* group (n=11), which suggests that awareness may play a facilitative role in learning.

The discussion so far has addressed studies from cognitive psychology with regard to attention and awareness in learning. Can such findings be generalized to SLA? Is *learning*, as it was operationalized in the above studies, analogous to the processes of language acquisition? Clearly, there are major differences between the input in these experiments and natural language. Syntactic structures are far more complex and ambiguous than ordered sequences and MAGs, which have no meaning or morphology.[4] There are also differences in the learning process itself.

---

[4] Schmidt (1994b) points out that there has been some psycholinguistic research using artificial languages, which have constituent structure and meaning, but this research has not addressed the issues of attention or awareness in learning.

SLA is largely an interactive process that involves the negotiation of meaning between interlocutors (Long, 1996). The type of learning demonstrated in the cognitive studies, by contrast, involved almost no feedback. Finally, differences in the findings between SRT and MAG studies suggest that differences in modality probably play a role in attentional demands. It follows that SLA might involve attentional demands that differ significantly from the stimuli used in cognitive studies. Nevertheless, structured sequences (especially MAGs) may provide a rough analogue of natural language because they reflect a complex underlying system of rules (Schmidt, 1994b). The investigation of how such rules are acquired with and without attention/awareness has provided insights into general cognitive processes that may be related to language acquisition but would have been difficult to capture using natural language as input. For example, differences between random and repeating or linear and hierarchical structures as well as previous linguistic knowledge, would have been difficult to control for using real language.

THE NOTICING HYPOTHESIS AND SLA

Much of Schmidt's work ties findings from cognitive psychology into SLA theory. As N. Ellis points out, "Schmidt is one of the few linguists who have adopted the conceptual and experimental rigours of experimental psychology in answering questions concerning the role of consciousness in L2 acquisition" (1994, p. 10). Reviewing the psychological literature on consciousness has led Schmidt to propose the *Noticing Hypothesis*, which states that "noticing is the necessary and sufficient condition for converting input into intake" (1990, p. 129)[5]. Since then, a considerable amount of research has addressed the issue of noticing in SLA. This section will briefly review some of that research. The discussion will mostly be limited to research designs that include some measure of noticing.

The noticing hypothesis seems to have been motivated by a seminal study by Schmidt and Frota (1986), which documents the role of noticing for a beginner learning Portuguese in Portugal over a period of 22 weeks. Extensive diary entries by the learner (Schmidt) were compared to tape-recorded interactions with native speakers to compare what had been noticed with what had been learned. Their findings question the assumption that language acquisition is a purely subconscious process (Krashen, 1982), since the learner clearly noticed some of the grammatical structures he seemed to have acquired. Different results were obtained in a similar study by Altman (1990, as cited in Schmidt, 1990), who monitored her own acquisition of Hebrew over a period of five years. She recorded her data in diaries, class notes, and underlined words in readings. These were compared to taped production tasks to see whether newly acquired vocabulary reflected what she had noticed in class as well as in naturalistic settings. Altman was unable to identify the source of half of the new verbs she had learned. She concluded that awareness was not necessary in learning vocabulary. Schmidt and Frota also admit that they were unable to trace much of what had been acquired to what had been noticed. Self reports are inherently subjective. Moreover, memory effects may play a role depending on the amount of time that passes before the diary entry is made. Nevertheless, first person accounts

---

[5] *Intake* refers to "a process that mediates between target language input and the learner's internalized set of rules" (Gass, 1988, p. 206). Intake may or may not lead to acquisition depending on how the input affects the learner's interlanguage.

seem to be the most valid method for assessing what is noticed. Other methods, such as posttests of target structures, allow only for inferences regarding the awareness of learners. Another source of validity in these studies is that they were longitudinal and included genuine communicative settings in and outside the classroom, which is quite rare in attentional studies.

One of the most influential attentional studies in SLA was conducted by VanPatten (1990), who investigated the notion of attention as a limited resource (Broadbent, 1958, as cited in Robinson, 1995; Wickens, 1980). More specifically, the study examined whether learners were able to consciously attend to both form and meaning when processing input[6]. Two hundred and two participants in university Spanish classes (levels I-III) were divided into four groups. All groups were presented with an audio recording of a 3-minute text and asked to listen for content. The control group did nothing else (content only). The other groups performed one of three additional tasks: (1) listening for the content word *inflación* (lexical); (2) listening for the definite article *la* (form); and (3) listening for the verb morpheme *-n* (morphology). Performance was assessed using a free written recall in English. Results showed that the *content only* and *lexical* groups significantly outperformed the *form* and *morphology* groups. This led VanPatten to conclude that it was difficult, especially for beginners, to notice content and form at the same time. Moreover, he postulated that learners would notice meaning before form, since their primary objective is to understand the propositional content of utterances.

VanPatten's findings have led SLA researchers to try and find ways to help learners focus on both form and meaning. One such way is input enhancement, which refers to the manipulation of certain aspects of the input (e.g., form) to make them more salient and thereby more noticeable to learners (Sharwood Smith, 1993). Typographical input enhancement usually entails italicizing, using boldface, or underlining in order to highlight the target structure. Alanen (1995) examined the role of typographical input enhancement and explicit rule representation on the acquisition of locative suffixes and consonant alternation in semi-artificial Finish. The input consisted of two passages with a picture and a Finish-English glossary of relevant words and forms. Participants were 36 university-level students with no prior knowledge of Finish. These were divided into a control group and three treatment groups according to different types of input, as follows: (1) italicizing the target structure (enhanced); (2) explicit rule presentation (rule); and (3) italicizing and explicit rule presentation (rule + enhanced). Performance was assessed with a pretest and a posttest, and think-aloud protocols were provided by participants as they read the passage. In terms of their performance on these tests, it was hypothesized that the treatment groups would fall into the following order: control < enhanced < rule < rule + enhanced. This pattern was only partially realized in the results as the quantitative analysis showed no significant difference between the enhanced and unenhanced input groups. However, think-aloud protocols revealed that learners who noticed the target forms learned more than those who did not.

A quasi-experimental study by White (1998) investigated the effects of typographical input enhancement and explicit instruction on sixth grade ESL students in a French elementary school. The study compared the performance of three treatment groups: (1) input enhancement + explicit instruction (n=27); (2) input enhancement (n=30); and (3) unenhanced input (n=29). The target structure was possessive determiners. Learning was assessed using an immediate and a delayed posttest after 5 weeks. In terms of performance, it was hypothesized that treatment groups would fall into the following order: group (1) > group (2) > group (3). Although the

---

[6] VanPatten's use of the term *conscious attention* seems to refer to attention + awareness (noticing).

accuracy ratio seemed to confirm this hypothesis, within-group variance cancelled out the expected between-group differences, suggesting that noticing did not have a significant effect. However, without introspective measures, there was no way of verifying if and what learners had actually noticed.

Stronger evidence for the facilitative role of noticing comes from a study by Jourdenais, Ota, Stauffer, Boyson, and Doughty (1995). Participants were 14 native speakers of English enrolled in an introductory level Spanish foreign language program. These were randomly assigned to two treatment groups: *Enhanced* and *Unenhanced*. The Enhanced group was given a Spanish passage in which preterit verbs were shadowed and imperfect verbs were bolded. Both verb forms were underlined and printed in a different font. The Unenhanced group received the same passage without any typographical modifications. The post-test consisted of a task in which participants had to write a picture-based narrative while providing think-aloud protocols. Results showed that the Enhanced group used the target structure more often than the Unenhanced group on both the think-aloud protocols and the written production task, suggesting that input enhancement made the target forms more noticeable. Moreover, subsequent production by the Enhanced group was more target-like than the Unenhanced group, suggesting that noticing facilitated acquisition.

Leow (2001) also used think-aloud protocols to examine how typographical input enhancement affects learners' noticing of the formal imperative in Spanish. Participants were 38 college-level students in a beginning level Spanish language program. The input consisted of a 242-word reading passage that was typographically enhanced for one group and left unenhanced for the other. Participants were asked to provide think-aloud protocols as they read the passage. The pretest, posttest, and delayed posttest consisted of a multiple-choice recognition task and a cloze test. A short-answer task was used to measure comprehension of the reading passage. Results showed that 33% (7 out of 21) of the enhanced group mentioned the target forms in their protocols as compared with only 12% (2 out of 17) in the unenhanced group. No statistically significant differences were found between the two groups for: (a) amount of reported noticing of the targeted form, (b) comprehension, and (c) intake as measured by recognition. However, significant correlations were found in both groups between noticing and recognition. Leow points out that the effects of typographical enhancement may have been diminished by the length of the input. When faced with a long reading passage, learners might be using more global noticing strategies in order to process the large amounts of input. This would probably shift attention toward meaning and away from form, since the former is more important for comprehension.

Leow's explanation seems to be supported by VanPatten's (1990) findings that attention to both form and meaning is difficult. However, the modality of the input in this case (written) differed from that in VanPatten's study (aural). Could modality differentially affect attention to meaning and form? Wong (2001) tried to address this question with a partial replication of VanPatten (1990). His variations included the addition of a written mode of input and using English (instead of Spanish). Participants were 85 low intermediate EFL students at a French Canadian university. Findings for the aural input mirrored those of VanPatten, since there was a significant decrease in performance when participants had to attend to both content and form. However, no significant difference was found when the input was written (which incidentally took less time to read than the aural input). Moreover, when processing both form and meaning, the listening task proved more difficult than the written task, suggesting once again that different modalities may impose different attentional demands.

A more innovative experimental design by Leow (1997, 2000) provides further evidence for the facilitative role of awareness in SLA. Leow (1997) used a crossword puzzle task as input that was designed to initially induce learner error. Eventual clues in the puzzle provided learners with the correct form, thereby increasing their chances of noticing the mismatch. Participants were 28 adult beginners in a Spanish foreign language course. The target structures were the irregular third-person singular and plural preterit forms of stem-changing *–ir* verbs. Participants were asked to provide think-aloud protocols as they solved the crossword puzzle. The post-exposure tasks consisted of a multiple-choice recognition and a written production task. Think-aloud protocols were analyzed using Allport's (1988) criteria for awareness, which divided participants into three groups: (1) *unaware* [+cognitive change, -meta-awareness, -metalinguistic description]; (2) *low aware* [+cognitive change, +meta-awareness, -metalinguistic description]; and *high aware* [+cognitive change, +meta-awareness, +metalinguistic description]. The study found that meta-awareness correlated with hypothesis testing and metalinguistic description. Moreover, the *high aware* group significantly outperformed the *unaware* and *low aware* groups on both post-exposure tasks.

Similar results were found in a subsequent study (Leow, 2000). Its research design and target population (n=32) were similar to the 1997 study but included three new features: (a) Participants were asked to answer two probe questions assessing their awareness immediately following the input and again after the post exposure task; (b) Participants were instructed to continue the think-aloud protocols as they solved the post-exposure tasks; and (c) The unaware participants were interviewed after the experiment and asked to explain their answers on the recognition task. The fact that these interviews took place three weeks after the experiment raises the issue of memory effects. Nevertheless, these additional features seem to increase the internal validity of claims regarding the awareness of learners. Again, results showed that participants who displayed evidence of awareness performed better on the post-exposure tasks than those classified as unaware.

In a similar experimental design, Rosa and O'Neill (1999) investigated the role of awareness in acquiring syntactic structures. Participants were 67 native English speakers enrolled in an intermediate Spanish course at an American university. These were randomly assigned to four groups based on the explicitness of instruction as follows: [-formal instruction, -directions to discover rules]; [+formal instruction, -directions to discover rules]; [-formal instruction, +directions to discover rules]; and [+formal instruction, +directions to discover rules]. The input consisted of a multiple-choice jigsaw puzzle that targeted counterfactual Spanish conditionals, which was followed by a multiple choice recognition task. Participants provided think-aloud protocols as they solved the puzzle to assess their level of awareness. Among other things, the study found that awareness seemed to increase learners' ability to recognize the syntactic structures on the post-test. There was also a strong correlation between awareness and intake.

DISCUSSION AND CONCLUSION

The review presented in this paper on attention and awareness in cognitive psychology and SLA is by no means exhaustive. Research in areas such as Corrective Feedback (Lightbown & Spada, 1990), Input Processing (VanPatten, 1996), and Interaction (Long, 1996), and Focus on Form (Long, 1996) clearly have important implications for the role of attention in SLA. Due to

limitations of space, the current review was confined to studies that included some measure of awareness, the key variable for testing Schmidt's noticing hypothesis.

The noticing hypothesis states that both attention and awareness are necessary for SLA. There appear to be several problems with this claim. One problem, alluded to earlier, relates to Schmidt's definition of attention (Schmidt, 1994a). Truscott (1998) points out that the definition of attention as alertness, orientation, and detection makes the claim that attention is necessary for learning seem rather obvious. He argues that since learning cannot possibly take place without detection, the claim that learning requires attention (if attention = detection) has "no empirical content" (p. 106). We may therefore conclude that attention is *by definition* necessary for SLA. To avoid misinterpretations, future research should perhaps avoid descriptions such as *unattended learning* altogether. What researchers in cognitive psychology have reported as unattended learning is perhaps more accurately described as *learning with low or residual attention*. Schmidt (1995) himself points out that it is more important to demonstrate that learning is enhanced by attention rather than to demonstrate that learning cannot take place without it. Most of the findings presented in this review support the claim that increased attention leads to more learning (Allport, Antonis, & Reynolds, 1972; Curran & Keele, 1993; Eich, 1984; Kellogg, 1980) but this may depend on the nature of the task since increased attention did not always lead to better performance (Cohen, Ivry, & Keele, 1990).

A second problem with the noticing hypothesis relates to evidence of implicit learning. A number of the studies reviewed in both cognitive psychology (Cohen et al., 1990; Curran & Keele, 1993; Hartman et al., 1989; Nissen & Bullemer, 1987) and SLA (Alanen, 1995; Jourdenais et al. 1995; Leow, 1997, 2000, 2001; Rosa & O'Neill, 1999; Schmidt & Frota, 1986; VanPatten, 1990) suggest a facilitative role for awareness in learning and SLA. However, there does not seem to be sufficient evidence to suggest that awareness is necessary for learning. This seems especially true in cognitive studies that showed evidence of implicit learning. Schmidt's (1995) contention that these studies did not demonstrate a complete absence of awareness seems justified in some cases (Hartman, et al., 1989) but not in others where participants were convincingly unaware (Cohen et al., 1990; Marcel, 1983; Nissen & Bullemer, 1987).

Tomlin and Villa (1994) also disagree with Schmidt on the necessity of awareness in learning. They argue for a dissociation between attention and awareness in learning, which brings them into conflict with the noticing hypothesis. However, according to Robinson (1995), "these different positions can be reconciled if the concept of noticing is defined to mean detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (p. 296). This conception of noticing is based on the idea that a certain threshold must be exceeded before input in short-term memory becomes part of awareness. Thus it becomes possible to learn without awareness but such learning is believed to have limited effects. Schmidt's recent writings have been more accepting of the possibility of limited learning without awareness, but he maintains that this type of learning is "of little potential benefit for language learning" (2001, p. 28). Yet, a moment's reflection on how some children learn a second language through interaction with peers and without any explicit instruction may lead us to conclude that implicit learning is not only possible but also highly effective for acquiring an L2. Children in these contexts must surely pay *attention* to the L2 input, but they are unlikely to be preoccupied with *noticing* (as defined by Schmidt) the linguistic surface structure of the input, nor are they necessarily *aware* of the complex grammatical rules underlying their interlanguage.

Despite its shortcomings, the noticing hypothesis has served to generate important theoretical and empirical debates in SLA. It has also provided an opportunity to integrate useful

concepts from cognitive psychology into SLA theory. But theories of attention/awareness in SLA have a long way to go in explaining what exactly gets noticed in the input (Truscott, 1998). Research is needed to determine what kind of interlanguage development results from noticing specific linguistic aspects. Do learners notice rules, exemplars, or both? Which stages of interlanguage devolvement are most appropriate for noticing specific structures? Any research design trying to answer these questions must include introspective data. Despite their subjectivity, introspective methods appear to be the only direct way of assessing what is noticed. Unfortunately, as Schmidt (1994a) points out, "the problem in applied linguistics has not been over reliance on first person reports and data, but an almost total neglect of them" (p. 22).

REFERENCES

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 259-302). Honolulu: University of Hawaii Press.

Allport, A. (1988). What concept of consciousness? In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 159-182). London: Clarendon Press.

Allport, A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology, 24,* 224-235.

Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition, 16,* 205-230.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17-30.

Chomsky, N. (1975). *Reflections on Language.* New York: Pantheon.

Curran T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189-202.

DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 27- 47.

Eich, E. (1984). Memory for unattended events: Remembering with and without awareness. *Memory & Cognition, 12,* 105-111.

Ellis, N. C. (2002). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition, 24,* 297-339.

Ellis, N. C. (1994). Implicit and explicit language learning: An overview. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 1-31). San Diego, CA: Academic Press.

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9,* 198-217.

Hartman, M., Knopman, D. S., & Nissen, M. J. (1989). Implicit learning of new verbal associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1070-1082.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior, 22,* 485-508.

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing? A think aloud protocol. In R. Schmidt (Ed.), *Attention and*

Awareness in Foreign Language Learning (pp. 183-217). Honolulu: University of Hawaii Press.

Kellogg, R. T. (1980). Is conscious attention necessary for long-term storage? *Journal of Experimental Psychology: Human Learning and Memory, 6*, 379-390.

Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition.* London: Pergamon.

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning, 47,* 467-505.

Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22,* 557-584.

Leow, R. P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84,* 496-509.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language acquisition. *Studies in Second Language Acquisition, 12,* 429-448.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197-237.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11,* 56-60.

Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1-32.

Norman, D. A. (1969). Memory while shadowing. *Quarterly Journal of Experimental Psychology, 21,* 85-93.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-420). New York: Academic Press.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25-42.

Posner M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *Foundations of Neuropsychology Series* (pp. 91-112). New York: Academic Press.

Posner, M. I., Walker J. A., Friedrich, F. J., & Rafal, R. D. (1987). How do the parietal lobes direct covert attention? *Neuropsychologia, 25,* 135-145.

Radford, A. (1997). *Syntactic theory and the structure of English: A minimalist approach.* New York: Cambridge University Press.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 77*, 317-327.

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

Robinson, P. (1995). Review article: Attention, memory, and the noticing hypothesis. *Language Learning, 45*, 283-331.

Rosa E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness: Another piece of the puzzle. *Studies in Second Language Acquisition, 21,* 511-556.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Schmidt, R. W. (1994a). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11,* 11-26.

Schmidt, R. W. (1994b). 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). San Diego, CA: Academic Press.

Schmidt, R. W. (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. 1-65). Honolulu: University of Hawaii Press.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). New York: Cambridge University Press.

Schmidt, R. W., & Frota S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to Learn: Conversation in Second Language Acquisition* (pp. 237-326). Rowley, MA: Newbury House.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*, 165-179.

Sperling, G. A. (1960). The information available in brief visual presentation. *Psychological Monographs, 74*, 1-27

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183-203.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research, 14,* (2) 103-135.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12,* 287-301.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

White, J. (1998). Getting the learner's attention. A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 85-114). Cambridge: Cambridge University Press.

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

Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and Performance* VIII (pp. 239-257). Hillsdale, NJ: Lawrence Erlbaum.

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23,* 345-368.

APPENDIX

The Entailments of Different Attentional Concepts Based on Schmidt (1994a)

Alertness

Orientation

Detection

Attention

Noticing

Awareness

Understanding