Selective Attention in Information Processing:  
A Review of Brain Science in SLA

Harumi OISHI

Abstract  Recent research in second language acquisition (SLA) has focused on whether and how language learners can best develop their target language acquisition skills. This paper discusses how selective attention helps successful language learners to process information in SLA. It also discusses the possibility of obtaining neuroscientific data in order to determine effective information processing in SLA. The SLA theory is based on human information processing. The well-known cocktail party effect (Cherry, 1953) describes the manner in which people pay attention to the linguistic information in everyday communication. It clearly has important implications with respect to the role of attention in SLA. The current review included some measures of attention that are the key variables for examining language learning. Most of the findings presented in this review support the claim that increased attention leads to more effective learning but this may depend on the nature of the task since increased attention does not always lead to better performance. The result of Oishi’s empirical study (2003) suggests that a considerably difficult task does not increase the blood flow so learners lose their attention while performing the task. The optimal result is achieved with an appropriate increase in the blood flow and attention, which facilitates language learning. The present study on SLA research addresses the neuroscientific perspective while learning English and sheds light on the findings from a number of areas of second language pedagogy and research.

1. Introduction

Recent research in second language acquisition (SLA) has focused on whether and how language learners can best develop their target language acquisition skills. In order to answer these questions, it is necessary to investigate the manner in which selective attention—which is deeply related to the noticing hypothesis in SLA literature (Schmidt, 1990)—functions in language learning. Due to the development of neuroimaging technology, SLA research is now considered as a part of neuroscientific investigation in order to clarify the manner in which successful learners process second language in their brain. Oishi (2003) conducted a new empirical study in language learning and teaching by using optical topography, with the conclusion that successful learners require optimal brain activation with the effective use of selective attention.

This paper discusses how selective attention helps successful language learners to process information in SLA. It also discusses the possibility of obtaining neuroscientific data in order to determine effective information processing in SLA.
First, basic human information processing is briefly addressed as a concept of first language acquisition. Second, the role of attention that facilitates SLA and the noticing hypothesis, which is deeply related to selective attention, is presented. Finally, the new field of second language study that makes use of neuroimaging technology is introduced in order to examine working of attention in the brain activity.

2. Cocktail party effect as the basis of SLA theory

In human information processing, the well-known cocktail party effect (Cherry, 1953) describes how people notice in a real communication. Although people simultaneously see and hear several things in daily life, they only select necessary information and ignore the remaining unnecessary information. Cherry (1953) suggests that if we focus on a conversation, our attention can be captured by a nearby conversation whenever an interesting message happens to come to our ears. The cocktail party effect describes the communicative ability to explain the “selective attention” of a single person amidst a mixture of conversations and background music, which ignores other conversations or sounds. This effect reveals the surprising abilities of human information processing mechanisms, which enable people to gather important information even in a noisy place.

Selective attention can be effective when we focus specially on the interesting information around us. For example, when talking with one of our friends in a crowded party, we can listen to and understand what the friend says even if the place is very noisy, and can simultaneously ignore the words of a nearby person. In another instance, when someone from the other side of the party room suddenly calls out our name, we notice that sound and respond to it immediately. This is known as passive attention.

These studies typically use dichotic listening tasks, in which participants wear headphones that provide them with a different aural message in each ear. Tasks are designed to heighten attention in one ear while ignoring it in the other. After exposure, an experiment is conducted on the amount of unattended information thus detected. For example, Cherry (1953) conducted an experiment in which subjects were asked to listen to two different messages from two speakers at a time and to pick up information from background noise and voice. He concluded that the information picked up was based on the characteristics of the sounds, the gender of the speaker, the speed of the speech, the direction from which the sound was coming, and the pitch.
3. Attention as a filter function and a limited capacity

Before discussing the role of attention in learning, the present section reviews some research on the nature of attention itself, specifically attention as a filtering process and as a limited capacity system (Sperling, 1960).

Robinson (1995) advocates two types of attention in cognitive psychology, both of which incorporate human memory. The filter theory, as the first category, addresses the notion of selection in attention. The underlying assumption in this theory is that there is only one channel of attention. Broadbent (1958) suggests a “filter model” by conducting dichotic listening experiments. The subjects were asked to hear and separate the different speech sounds that were simultaneously presented in both ears. The result of this experiment suggests that “our mind can be conceived as a radio receiving many channels at once”: the brain separates incoming sounds into channels based on physical characteristics and submits only certain subsignals for semantic analysis. It is also possible that there exists a type of listening filter in the human brain that selects the information to which an individual must pay attention from among the mixture of information that she or he perceives. This theory is based on the “bottleneck” model (Broadbent, 1958). In this model, input is registered, selected, detected, and entered into short-term memory by the senses. This selection model assumes that only form can be noticed at the early selection stage, and that meaning is analyzed once the input is passed through this filter. According to other selection models, both form and meaning are analyzed prior to detection; however, this would involve substantial mental processing of the input even before noticing occurs. Finally, there are also filter models in which selection does not occur till the input enters short-term memory. In these models, while the selected input is rehearsed, the unselected input is forgotten.

The other kind of attention is based on capacity theories. These theories do not assume that the stimuli are necessarily competing for attentional resources. Instead of describing attention as a filter, they use the metaphor of a spotlight, which can be a narrowed focal attention as well as a broadened global attention. The key factors in capacity theories are the voluntariness of attention and the attentional demands of the task itself. Wickens (1980) argues that attentional demands become more difficult when simultaneous tasks draw on the same attentional pools, such as the effort made 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, when attentional demands differ, parallel processing is possible; the quality of performance, however, may suffer as a result of this. For example, most people can simultaneously drive and talk on their cell phones; however, many countries, including Japan, outlaw talking on the cell phone while driving because it seems to distract the focal attention.
from driving. Having pointed out that individuals may differ with respect to memory and parallel processing ability, Wickens (1980) suggests that such variables need to be accounted for in attentional studies. Most theories assume that attention plays a positive role in processing language input because it involves detection and selection. However, attention may also be viewed as an inhibiting mechanism that prevents linguistic 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.” Lack of attention prevents learners from focusing their attention on given information. In such cases, individuals are unable to inhibit a potentially attention-capturing interference. This suggests that inhibition plays an important role in attentional processes. These notions also form the basis of the use of attention in the development of SLA skills.

4. **Attention and Knowledge in SLA**

According to Tomlin and Villa (1994), there are four conceptions of attention in SLA. One is that attention is a limited capacity system. In other words, an idea may be the result of a brain mechanism and conveyed through the sensory system, with an overwhelming linguistic information at any given time; it seems impossible to process them all at once. The limitations of attention refer not only to the amount of attention that may be given to a single message but also the amount of information that may be simultaneously attended to. This is related to the second conception of attention, the process of selection of information. The overwhelming amounts of incoming information force the attentional system to be selective. The third conception is attention under consciousness. It involves controlled rather than the automatic processing of information and is based on the assumption that controlled processing requires more effort and time to process information, and a higher degree of attention than does automatic processing. The conception of controlled processing of two simultaneous tasks is possible develops the fourth conception of attention which is that it must involve a process of coordination among competing linguistic stimuli and responses.

Therefore, if the processing is automatic with low attention and low consciousness, a person can perform two tasks at the same time; however, if both of them are control processing with high attention and high consciousness, it becomes more difficult to process two tasks simultaneously. Paradis (1994) suggested two types of knowledge. One is implicit knowledge, which is acquired without consciousness. Even after acquisition, this type of knowledge cannot be accessed by conscious memory and cannot be put to use spontaneously without conscious control. The second type is
explicit knowledge, of which the learner is conscious and can access on demand like the metalinguistic knowledge of grammar. Some researchers state that interaction between the two types of knowledge is impossible (Krashen, 1982; Paradis, 1994; Truscott, 1998). On the other hand, others claim that such interaction is possible on a linear learning processing, and that explicit knowledge can be converted into implicit knowledge. This knowledge plays a crucial role in the noticing hypothesis (Carr & Curran, 1994; Ellis, 1994; Schmidt, 2001; Robinson, 1995). In any case, however, even if the two processes are related to each other, it is important to note that although attention is necessary for learning, such explicit learning does not necessarily result in explicit knowledge.

Early experiments investigating the role of attention in learning tended to be selective attention studies, which required participants to attend to certain information and ignore the other information. According to Schmidt (1995), studies that claim to provide evidence for learning without attention only succeed in revealing that less attention leads to less learning.

5. Selective attention as a metacognitive strategy

There are strong relationships among reading strategies, metacognitive awareness, and reading proficiency. Selective attention here is considered as a metacognitive strategy in language learning. O’Malley and Chamot (1990) recommend the effective use of some metacognitive strategies that promote learning, including selective attention, planning, monitoring, and evaluation. Selective attention in language learning promotes the understanding of new information, choosing and integrating out of a student’s background knowledge. Learning is promoted when a student uses the metacognitive strategies effectively. Successful readers are likely to employ more metacognitive strategies than less successful readers; moreover, better learners possess enhanced metacognitive strategies that lead to greater reading ability and proficiency (Baker & Brown, 1984; Garner, 1987; Pressley & Afflerbach, 1995).

Many researchers have found that in general, proficient readers first look for an overview of a text before reading the text and employ context clues such as titles, subheadings, and charts or diagrams; they also look for a gist while reading, pay greater selective attention to important information, and attempt to relate important points in a text to one another in order to well understand the text as a whole. According to Carr and Curran (1994), “if you are conscious of something (language), then you are attending to it…and if you are attending to something (language), then you are conscious of it.” In learning a second language, as study progresses, learners accumulate knowledge and hence increase the capacity of their attention. The capacity changes with task difficulty, teaching method, and contents of the task, as
mentioned earlier.

6. Selective attention as a brain function

According to Ellis (2002), “we are now at a stage at which there are important connections between SLA theory and the neuroscience of learning and memory.” The concept of attention in language learning has become especially important, because the role of attention is important in many aspects of SLA, such as input, processing, development, and instruction; moreover, attention is a central issue in cognitive psychology (Case, 1985). Posner (1995) explains the function of selective attention in the human information processing model. He states that the role of selective attention is to preserve the information processing system from information overload. The brain will overflow with information without a filtering process. The intellectual development of humans was constrained by the biological maturation of the brain. According to Baldwin (1895), the brain has a limited capacity in terms of attention span as a measure of mental capacity, which a person could attend to at a time. He believes that the attentional capacity of an individual is biologically determined. Neuroanatomical evidence would support the role of attention (Posner & Petersen, 1990). Monitoring electrical brain activity or brain blood flow by using neuroimaging technology, such as event-related potentials and positron emission tomography (PET), has enabled researchers to relate attentional networks to the linguistic area of the brain (Posner & Petersen, 1990).

We receive information through the filtering process. However, the filtering process should completely cut off the unattended brain activation, because the system of information processing in brain mechanism might otherwise miss important signals. The theories of selective attention explain this complicated issue by assuming that only information selected is efficiently processed. This selected information generates consciousness in language learning.

Baars (1983) gives examples of contrasting conscious and unconscious processing. Newman and Baars (1993) have developed a neural attentional model that indicates several distinctions of consciousness. This model demonstrates that consciousness requires an activation process beyond that which is generated by the stimulus itself. The traditional metaphor known as “searchlight” explains consciousness as a searchlight for providing the necessary signals. This searchlight generates the perception or idea to enter into awareness.

Taking filtering processes into account, the neuroscientist Crick (1984) pointed out the functions of a thalamic in the brain searchlight. Baars (1983; 1988) and Newman and Baars (1993) have extended this metaphor of existence of thalamic in the brain to brain activities. This extended reticular–thalamic activating system (Baars, 1988) is
postulated to provide the supplementary activation that is required to activate the modular processes of input. The function of an attentional system is to select several potential inputs that will become conscious at any given time. The reticular nucleus of the thalamus (nRt) in the brain functions as a gate traffic from sensory surfaces to the cortex.

The SLA theory is based on such human information processing theories. Oishi and Kinoshita (2003) compared cortical activation patterns in first and second languages in language-related areas such as the Wernicke’s area, primary auditory area, supramarginal gyrus, and the angular gyrus, while language learners listened to and performed first language (L1) and second language (L2) comprehension tasks by a new noninvasive technique known as optical topography. The result showed that the amount of increased blood flow in the L2 task was more than that in the L1 task. This suggests that an easy task requires less attention and hence a lesser amount of blood flow than that required in a difficult task.

Likewise, SLA is required to only select necessary information from among many items to generate effective performance. Good learners tend to select important information with speed, fluidity, and accuracy. In SLA, selective attention plays an effective role in language comprehension. Learners obtain the optimal input when they pass information through the filtering process in the brain mechanism. An examination of the brain activation using a recently developing neuroimaging device helps to clarify the mechanisms of SLA.

7. Attention functions in noticing hypothesis

The noticing hypothesis, which discusses the role of noticing for a beginner, has been propounded by Schmidt and Frota (1986). According to Schmidt (1990), “intake is that part of the input that the learner notices.” He also identifies the three aspects of consciousness involved in language learning: awareness, intention, and knowledge. The first aspect, consciousness as awareness, embraces noticing. Although these definitions of noticing differ in terms of degree or level of awareness that are deemed necessary for SLA, all the researchers agree on the importance of noticing in SLA. According to Schmidt (1993), second language learners need to not only understand the input information but also notice “whatever features of the input are relevant for the target system.”

There are both strong and weak versions of the noticing hypothesis. The strong version insists that noticing is necessary and sufficient for SLA; the weak version states that noticing is a necessary but not sufficient condition for SLA. Noticing, however, is considered to play a significant role in SLA in both the strong and weak versions. According to Schmidt (1995), “the noticing hypothesis states that what
learners notice in input is what becomes the intake for learning.” He also states that whether a learner deliberately attends to a linguistic form in the input or whether it is noticed purely unintentionally, if it is noticed, it turns to intake; such noticing is a necessary condition for SLA.

In Gass and Selinker’s (2001) view, noticing or selective attention is “at the heart of the interaction hypothesis” and is one of the crucial mechanisms in the negotiation process in SLA. According to Ellis (1994), “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.” A review of the psychological literature on consciousness led Schmidt (1990) to propose the noticing hypothesis, which states that “noticing is the necessary and sufficient condition for converting input into intake.” Since then, a considerable amount of research has addressed the issue of noticing in SLA.

VanPatten (1990) studied influential attention in SLA. He investigated the notion of attention as a limited resource (Broadbent, 1958; Wickens, 1980) and concluded that it was difficult, especially for beginners, to simultaneously notice both content and form. In addition, he assumed that learners would notice meaning before noticing form, since their main objective would be to understand the gist of the message. VanPatten’s findings suggest that SLA researchers should find ways to help learners focus on both form and meaning. This leads to input enhancement, which refers to manipulating the input in order to aid the understanding of learners’ (Sharwood-Smith, 1993).

8. How to utilize optical topography in SLA studies

How do people acquire language skills? Chomsky (1957) first advocated universal grammar (UG) to answer this question. UG is a revolutionary concept in linguistics and cognitive science (Chomsky, 2000). In order to understand the mechanism of language acquisition, it is important to consider how language is related to the several aspects of the mind and brain activities. Probably inspired by Chomsky’s modular approach to the study of the mind (Chomsky, 1980), Fodor (1983) postulated “the modularity of mind,” which attributed language to an input system as a module like the sensory system. However, according to Chomsky (1980), it is too narrow to regard only the “language module” as an input system; in 1986, he claimed that language acquisition is rather a “central system.”

With regard to the modularity of language and memory, there is experimental evidence. Normal speech comprehension and speech production are possible only in a conscious state. Dennett (1978) noted, “if one supposes…one’s thinking is one’s ‘stream of consciousness.’” According to the hypothesis of functional modules,
individual modules constitute a system as a whole. If the modules of phonology, semantics, and syntax conceivably exist in the language system, then information flow among these modules as well as the roles of their interactions should be clarified. A critical question is whether these modules correspond to the distinct areas of the brain (Sakai et al, 2000).

The first step toward investigating how language functions in the human brain is identifying specialized cortical areas in which language processing takes place. Lesion studies of the language areas of the brain have suggested three candidate areas for language function. In Brodmann’s language map, Wernicke’s area (22), AG: SMG the angular gyrus (39), the supramarginal gyrus (40), and Broca’s area (44, 45) are considered to be the regions of language function. However, language function is often too broad to be specified in particular areas. The individual roles of these areas are still being debated, and the exact correlation between cortical language areas and the subcomponents of the linguistic system have not been established yet (Sakai et al, 2000).

The recently developed functional imaging techniques, however, address these critical issues. Such imaging techniques should be combined with detailed anatomical studies in order to contribute to an understanding of both the structural and functional bases of language information processing.

The fundamentals of the optical topography system are based on such language phenomena. The system uses near infrared light to measure the changes in oxyhemoglobin and deoxyhemoglobin as well as the total amount of blood changes in various areas of the cortical surface as images. When a specific area of the brain is activated, the localized blood flow in that area quickly increases. Thus, it can be determined where and how active the specific areas of the brain are by continuously monitoring blood hemoglobin levels, while having the examinee perform some language tasks. It takes only 0.1 second to perform one cycle of measurement, calculation, and display; thus, this system provides a “real-time” measurement of the brain’s activity. There are several advantages of using optical topography over current measurement methods and functional mapping techniques such as fMRI and PET. It is noninvasive and can be used in a variety of conditions with minimal restriction on the examinee. Measurements can be made in more natural conditions like language classroom, giving more freedom in choosing from among varieties of language tasks, including listening, speaking, reading, and writing. There are no scanning noises to interfere with processing language tasks. It does not require a head constraint to ensure that participants sit in a usual study position throughout the experiment. However, the one disadvantage of this system is that it restricts measurement from the cortical surface; nevertheless, this system provides a new dimension for mapping the human brain and cognitive functions.
The infrared light passes through the skull and reaches the cerebral cortex. It penetrates to a depth of approximately 30 mm and is scattered through hemoglobin in the blood. The light is partially reflected back through the scalp. The reflected light on the scalp contains the information on the cerebral cortex. A 30-mm distance from the original illuminated point is easy to detect. The reflected light is then sent through optical fibers and received with a highly sensitive avalanche photodiode detector. The brain activity is characterized by the firing of neurons. This results in active energy and causes a secondary increase in the blood flow to supply oxygen to the brain tissues. Thus, the change in the concentration of hemoglobin is an important index for knowing the brain activity (as stated by Hitachi Co.)

9. Conclusion

The review presented in this paper discusses the nature of attention in SLA, such as Schmidt’s noticing hypothesis. The theory of attention is based on human information processing. The well-known cocktail party effect (Cherry, 1953) describes the manner in which people notice in everyday communication. Cherry (1953) suggests that if we focus on a conversation, our attention can be captured by a nearby interesting message.

The SLA theory is based on human information processing. It clearly has important implications with respect to the role of attention in SLA. The current review included some measures of attention that are the key variables for examining language learning. According to the noticing hypothesis, attention is necessary for SLA. Schmidt (1995) 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 effective 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 does not always lead to better performance (Cohen, Ivry, & Keele, 1990).

Currently, research is needed to determine the kind of language development that results from noticing aspects in terms of objective data as neurolinguistic evidence in addition to cognitive evidence. Despite their subjectivity, introspective methods is the only direct way of assessing what the learners notice. Schmidt (1994), however, 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.”

Neuroanatomical evidence would support the description of the role of attention (Posner & Petersen, 1990). Monitoring the electrical brain activity or brain blood flow by using neuroimaging technology, has enabled researchers to relate attentional
networks to certain areas of the brain. The result of Oishi’s empirical study (2003) suggests that a considerably difficult task does not increase the blood flow so learners lose their attention while performing the task. The optimal result is achieved with an appropriate increase in the blood flow and attention, which facilitates language learning.

Accordingly, the functions of attention are associated with the increased activity of the brain (Posner & Petersen, 1990; Posner & Rothbart, 1992). 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.”

This study hypothesizes that neuroscience can provide plausible mechanisms for understanding how learners internalize linguistic systems available in a comprehensible input and intake. The present study on SLA research addresses the neuroscientific perspective while listening to and reading English and sheds light on the findings from a number of areas of second language pedagogy and research.

Accepted May, 9, 2007

REFERENCES


Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two


Linguistics, 11, pp.129-158
-----------(1994). Deconstructing consciousness in search of useful definitions for applied linguistics. AILA Review, 11, pp.11-26