Input Processing in Second Language Acquisition: A Discussion of Four Input Processing Models

Yayun Anny Sun

Teachers College, Columbia University

ABSTRACT

The importance of input has been a widely recognized concept in the field of second language acquisition. Much research has been conducted to examine how input is processed, the various facilitative attributes of input, and the effectiveness of pedagogies that directly manipulate input. Among these domains of input-related research, it is critical to first understand the very nature of input-processing. Hence, in this paper, four different models of input-processing are examined and compared side-by-side. The discussion aims to disambiguate discrepancies in terminologies, identify common emphases on gap-noticing and cognitive-comparison, and suggests the need for further research on the role of attention/consciousness in input-processing.

INTRODUCTION

Input is one of the most important elements in the process of second language acquisition (SLA). As Gass (1997) points out, second language (L2) learning simply cannot take place without input of some sort. This statement has been generally supported by researchers in the field regardless of one’s theoretical approach. Building upon this understanding, specific issues have been actively debated, such as: (1) how input is processed during SLA and how it is incorporated into a learner’s developing interlanguage (IL) systems (Carroll, 1999, 2000; Chaudron, 1985; Gass, 1997; Krashen, 1982; Sharwood Smith, 1986, 1993; VanPatten, 1996, 2002); (2) the amount of input that is necessary to enable acquisition (Ellis, 2002; Krashen, 1982; White, 1989); (3) the various attributes of input and how they may facilitate or hinder acquisition (e.g., frequency, saliency, and transparency); and (4) instructional methods that may enhance input to promote acquisition (e.g., various types of input enhancement, recasts, and processing instruction). The above four domains of inquiry have led to a plethora of studies. However, before one can logically approach the issues related to application, as in the latter three debates, there must first be an overall understanding of how input is in fact incorporated into the interlanguage grammar, as highlighted in the first issue. Decades of discussion on input processing have produced useful insights but also diverging terminologies and models. It is pertinent to understand them holistically in order to further the discussion in a more organized and efficient manner. Hence, in this paper, I will first seek to clarify the definitions of key terms. Next, I will explore four input processing models, proposed by

---

1 Yayun Anny Sun is a doctoral student in Applied Linguistics at Teachers College, Columbia University. Her research interests include second language acquisition, especially input processing and instructed SLA, as well as teacher education and classroom discourse. She is also a full-time ESL instructor at Borough of Manhattan Community College, CUNY. Correspondence should be sent to Yayun Anny Sun, Teachers College, 525 W. 120th Street, Box 66, New York City, NY 10027, E-mail: yas8@columbia.edu
Chaudron (1985), Sharwood Smith (1986), Gass (1997), and Carroll (1999, 2000), extrapolate their commonalities and discrepancies, and identify key issues that demand further investigation.

CURRENT CONSENSUS ON INPUT

Corder (1967) made a significant observation in his seminal paper about how input is perceived in the process of L2 acquisition. His insight later became one of the cornerstones in input-related research. He discussed the notion of *intake*:

The simple fact of presenting a certain linguistic form to a learner in the classroom does not necessarily qualify it for the status of input, for the reason that input is “what goes in” not what is available for going in, and we may reasonably suppose that it is the learner who controls this input, or more properly his intake. (p. 165)

The fact that not all the available data in the learner’s environment can be absorbed and used in building the learner’s IL grammar presents one conundrum, and the condition that would enable the conversion of input into intake has been a central point of research. Corder’s comment also shifted the way SLA researchers perceived input: from a strictly external phenomenon to the interface between the external stimuli and learners’ internal systems. Discussions on learners’ developmental readiness, teachability, and other cognitive factors thus came to the fore (e.g., Doughty, 2001; Pienemann, 1989). The common consensus in the field of SLA is that what input learners are actually able to use for developmental purposes will depend on their current state of knowledge. Following this acknowledgement, however, it remains unclear exactly what mechanisms and subprocesses are responsible for the input-to-intake conversion. The models described below will provide some insights into this question.

Another basic understanding for SLA researchers is that, as input is converted into intake, learners make use of this material for dual purposes, namely, comprehension and acquisition. Drawing this distinction is important for both theory-making and empirical investigations (Faerch & Kasper, 1980; Krashen, 1982; Sharwood Smith, 1986; Swain, 1985; VanPatten, 1996). Learners have the natural inclination to decode linguistic input for meaning to achieve successful communication. But the type of intake derived from processing-for-meaning is not equivalent or sufficient to that which is needed for acquisition, which entails the creation of new or revised mental structures. Swain’s (1985) study of a French immersion program revealed that, based on communicative and comprehensible input alone, learners may achieve native-like proficiency in their comprehension. But their proficiency and accuracy in production lags behind that of native-speakers despite years of exposure. Swain’s study provides support that comprehensible input does not necessarily lead to acquisition. At the same time, it is also true that, without comprehensible input, learners would not be able to make the necessary form-meaning connection for acquisition to occur; the reasons for this will become apparent in the discussion of the models below (Krashen, 1982, 1985; VanPatten & Cadierno, 1993). This distinction between processing for comprehension and acquisition is another concept that is commonly accepted in the field of SLA research. The models below focus on how input is processed differently for comprehension and acquisition, and they lay out additional processes that need to occur beyond comprehension to trigger acquisition.

The two dichotomies just discussed (i.e., input/intake and comprehension/acquisition) serve as a starting point for the models reviewed below. Beyond the general consensus, the views on
input processing and L2 acquisition begin to diverge. The discussion will now turn to the models themselves to highlight their commonalities and differences.

MODELS AND DISCUSSION

Responding to Corder’s (1967) insight and recognizing that the use of the term input has often been inconsistent in the field of SLA, Chaudron (1985) attempted to disambiguate and reconceptualize input processing. Chaudron’s model is presented first because it provides a stepwise framework that is paralleled by many of the later models. By a close examination of this model, further questions can be generated and clarification sought in the other models. The four models, proposed by Chaudron (1985), Sharwood Smith (1986), Gass (1997), and Carroll (1999, 2000), respectively, are chosen because they all subscribe to a modular, nativist view on acquisition, in which linguistic processing is considered unique from other types of learning, and is served by a linguistic-specific processor. This is different from the connectionist approach, in which input is basically regarded as tokens for cue-weights and frequency-counting, and from which the statistically-based constructions arise (Ellis, 2002). The connectionist frameworks deserve a separate review all their own and thus will not be included here.

Returning to the discussion of Chaudron’s (1985) model, essentially, it consists of three intake stages. They are: (1) the preliminary intake (i.e., the perception of input), (2) the subsequent stage of recoding and encoding of the semantic information into long-term memory, and (3) final intake (i.e., where learners fully integrate and incorporate the linguistic information in the input into their developing grammars). Two separate bodies of research were incorporated to capture these sub-processes. For the preliminary and subsequent intake stages, Chaudron drew on a first language (L1) information processing model put forth by Massaro (1975). According to the general view on information processing, a neural-based and bottom-up signal processing takes place initially, where auditory feature detectors receive speech (or visual) signals as neural impulses and analyze them according to constraints evolved in the detectors. The analyzed input is then stored in short-term storage, where linguistic rules and other knowledge systems are called upon from long-term memory to interpret this filtered signal and to synthesize it into phoneme and word-strings. Then, as the surface structures fade in short-term memory, a more abstract representation of the speech, through rehearsal and recoding, is retained in long-term memory. During this input-comprehending phase, processing operates both in a bottom-up and top-down fashion, that is, there is a continuous interaction and exchange of information taking place in working memory, between the feature-analysis and predictions made based on acquired rule systems. This sequence completes the comprehension aspect of input processing.

After comprehension takes place in the first two intake stages, learners may proceed to the third stage where their IL grammar is restructured and developed. Chaudron subscribed to Faerch and Kasper’s (1980) and Krashen’s (1982) complementary views. By having comprehended the input using both their current L2 competence and extra-linguistic knowledge, learners may notice the gap between their current IL grammar (i.e., i) and the i + 1 presented in the input, which would become the material that triggers their next step of development (Krashen, 1982, 1985). It is not clear in this model exactly how certain aspects of linguistic structures are noticed and selected and how new rules emerge, but learners’ readiness and a natural order hypothesis have been suggested for a partial answer (Pienemann, 1989). Then, once the gap has been noticed, the learner’s innate language acquisition device (LAD) subsequently uses these new materials to formulate IL rules and perform hypotheses-testing. After the rule has been initially formulated, learners’ output
production and the feedback they receive would then serve as a platform to test, confirm, or revise the rules. With sufficient testing, reformulation, and confirmation, the new rule is then incorporated into learners’ IL grammar.

With the three-stage conceptualization, Chaudron’s (1985) model provides an overview of input-processing and a point of departure for further discussion. Several components in the model demand further explanation. For one, interestingly, Chaudron evoked the notions of automatic and controlled processing to describe the quality of processing during the preliminary and secondary stages (Shiffrin & Schneider, 1984). By definition, controlled processing is malleable and open to explicit training (DeKeyser, 2001). The implication then is that, other than the hardwired and neural-based mechanisms, initial processing may be open to a learner’s conscious manipulation, which then can be made automatic through repetition. This implication begs the question of whether learners would be able to consciously retrieve phonemic information from long term memory to interpret incoming signals. If that is the case, this point would bear substantial theoretical as well as instructional significance. However, other models of input processing do not seem to support this, as will become more apparent in the later discussions. Or at least, Chaudron’s model needs to provide further specification to this point. Another point of ambiguity that is more crucial and problematic is the notion of gap-noticing. It is presented in this model as the central processing that needs to take place before a structure can be acquired. But it is not clear exactly when and where this operation takes place. Is it an encapsulated sub-process that takes place in the innate LAD? Does introspection or awareness of this gap-noticing play any role in facilitating acquisition? These questions have been left open in Chaudron’s model of input processing. Seeking clarification for these questions from a complementary view, I now turn to Sharwood Smith’s (1986) model.

In his model, Sharwood Smith (1986) focused and elaborated on the acquisitional aspect of input processing, which is closer to Chaudron’s (1985) notion of final intake, and bypassed the initial signal processing component in his explanation. He agreed with Chaudron’s view of hypothesis-testing, which he termed the making of mental comparisons, and he asserted that this mental comparison indeed operates within the LAD. His five-stage acquisitional procedure starts out with learners making comparisons between their semantic representations (derived purely from current linguistic competence) and the total meaning representations (derived from competence and extra-linguistic and world knowledge). In the second stage, learners adjust their semantic representations as they compare the two sets of representations. Third, learners generate a surface structure from the adjusted semantic representation, using rules in their current grammar. Fourth, learners compare the original surface structure with the new surface structure and note any discrepancy. Finally, learners restructure their current competence system so that the adjusted semantic representation may be derived from the surface structures encountered in the future. During this five-stage operation, Universal Grammar (UG) and learners’ L1 knowledge may also come into play and mediate the entire process.

Similar to Chaudron’s (1985) model, Sharwood Smith (1986) emphasized the derivation of meaning before acquisition. Without first comprehending the messages, learners would not be able to proceed with the first step of comparing semantic representations (stage 1). In this sense, Sharwood Smith, like Krashen (1982), generally promoted the use of rich and complex input instead of simplified input. This is so that learners can utilize the additional extra-linguistic cues, available in the linguistic environment, to derive semantic representations when their current linguistic competence cannot support a full understanding. However, comprehensible input alone is
not sufficient, and acquisition may still break down for a number of reasons, which will be explicated below.

According to this model of input processing, there exist several junctures where processing for acquisition may break down. First, if there is not a noticeable gap between the two semantic representations (i.e., no glitches in comprehension), even if differences may indeed exist, learners would not attend to it. Underlying this phenomenon are the issues of depth of analysis (i.e., whether learners stop short at comprehension or continue to deeper analysis and actually process the form) and attention (i.e., whether learners would attend to the gap, with conscious awareness or not; Schmidt, 1990; Schmidt & Frota, 1986; Tomlin & Villa, 1994). The second possible breakdown in acquisition is that there is also the danger of learners conjuring an incorrect overall representation based on extra-linguistic information, which then would lead to inaccurate mapping. This is one place where repeated exposure becomes critical for fine-tuning the representations and hypotheses. Finally, even if learners are aware of the discrepancy between the two semantic representations in stages 1 and 2, there is no guarantee that they would proceed to the third stage, where a new surface structure is generated to better fit the adjusted representation. In fact, this suggests an irony in the model. If the current competence is insufficient to generate a correct semantic representation in the first stage (which implies the presence of incorrect rules or the complete absence of rules), how are learners able to generate a new surface structure from their incomplete competence for further comparison (stages 3 and 4)? What mechanisms, if not learners’ competence, would induce this generation of new structures? This last logical problem was not addressed directly by Sharwood Smith (1986) in his model and remains a conundrum. However, the issues related to noticeable-gap, depth of analysis, and attention are further addressed in Gass’s (1997) model, to which I will now turn.

Gass’s (1997) framework of SLA includes a similar sequence as Chaudron’s (1985) and Sharwood Smith’s (1986) models with stages of apperceived input, comprehended input, intake, integration, and output. Gass made a finer distinction between apperceived input, comprehended input, and intake, in which the first stage of apperceived input, in Gass’s conception, is not a bias-free processing. A certain level of recognition and selection has already taken place as a result of attention, in the similar sense as Tomlin and Villa’s (1994) notion of orientation. Paralleling Chaudron’s and Sharwood Smith’s models but going beyond, Gass particularly stressed the importance of negotiated interaction during input processing and acquisition. Failure in communicative interaction pushes learners to negotiate for meaning. Through the act of clarification and elaboration for comprehension, learners then receive additional and usable input, and their attention may be drawn to specific problematic features in the L2. Consequently, interaction increases the chance for learners to make mental comparisons between their IL and the L2, in the same sense as proposed by Sharwood Smith. Through negotiated interaction, the input is enhanced in three ways. First, it is made more comprehensible, which is a prerequisite of IL development. Second, problematic forms that impede comprehension are highlighted and forced to be processed to achieve successful communication. Third, through negotiation, learners receive both positive and negative feedback that are juxtaposed immediately to the problematic form, and the close proximity facilitates hypothesis-testing and revision (Doughty, 2001). In light of its three-fold effects on acquisition, the interaction component of Gass’s model really should be regarded as a facilitator of learning, not a mechanism for learning.

As a midpoint summary, the models reviewed so far, proposed by Chaudron (1985), Sharwood Smith (1986), and Gass (1997), converge on the necessity of comprehensible input (or comprehended input, in Gass’s term). Learners must be able to decode enough of the input to
formulate a conceptual representation, through which linguistic structures can be called upon from current competence and be compared with the external and apperceived structure. Paradoxically, but perhaps not incompatibly, there must also be incomprehensible input—some extra bits of linguistic forms that cause a mental jolt in processing. Had everything in the input been completely understood, learners would generally feel no need to attend to forms, and acquisition of missing structures would not occur. In other words, because of the incomprehensibility of the input, learners’ attention is drawn to the specific structure. Then cognitive comparison between IL representation and external representation would take place, which would eventually lead to acquisition (Gass, 1997; VanPatten, 1990; White, 1987). Gass’s interaction model aims to create exactly such moments, as already described above.

An example of this interplay between comprehensible and incomprehensible input is the Processing Instruction approach created by VanPatten and his collaborators (Slobin, 1985; VanPatten, 1996, 2002; VanPatten & Cadierno, 1993a, 1993b). With this approach, they try to induce failure and generate the need for learners to divert from their default processing strategies to attend to the specific form that requires acquisition. Although it is not certain at this point whether learners’ innate processing strategies themselves can really be altered through this method, it has at least included both the comprehensible and incomprehensible input which need to work in tandem for the development of IL grammar. For that purpose, the conception of Processing Instruction adheres to the key views presented in the above models, and its instructional results have generally been successful (e.g., Benati, 2005; Cheng, 2004; Morgan-Short & Bowden, 2006; VanPatten & Cadierno, 1993). As can be seen, the thoughts on comprehensible and incomprehensible input in fact converge.

The last model reviewed here is Carroll’s (1999, 2000) Autonomous Induction Theory. It provides a contrast to the above views on the primacy of comprehension and attention. Carroll’s model is adapted from Jackendoff’s (1987) modularity model and Holland, Holyoak, Nisbett, and Thagard’s (1986) induction model. Basically, linguistic faculty is comprised of a chain of representations, with the lowest level interacting with physical stimuli, and the highest with conceptual representations (cf. Carroll, 1999). Two types of processors are at work for each level of representation: the integrative processor combines smaller representations into larger units, and the correspondence processor is responsible for moving the representations from one level to the next (e.g., from the acoustic level to the phonological level). Together, the two processors form an encapsulated and sequential module. At each tier of the encapsulated modules, the representations are categorized and combined according to UG-based or long-term memory-based rules. This procedure outlines how input is processed for parsing.

Carroll (1999, 2000) makes a clear distinction between processing for parsing and for acquisition. It is exactly when the parsers fail that the acquisitional mechanisms are triggered—a view that is somewhat aligned with the notion of incomprehensible input. But instead of using a very general notion of noticing the gap and cognitive comparison, Carroll spells out the sequence of restructuring and enhances the understanding on this somewhat vague area. Namely, during successful parsing, rules are activated in each processor to categorize and combine representations. Failures occur when the rules are inadequate or missing. Consequently, the rule that comes closest to successfully parse the specific unit would be selected and would undergo the most economical and incremental revision. This process is repeated until parsing succeeds or is at least passable at that given level. This procedure explains the process of acquisition, where the exact trigger for acquisition is parsing failure resulting from incomprehensible input.
Continuing with the discussion of comprehensible and incomprehensible input, Carroll (2000) contradicts the way Gass (1997) conceptualizes and sequences input-processing in her model. Gass conceives of intake as a subset of comprehended input. However, according to Carroll’s logic, comprehension involves the extraction of meaning to form conceptual representations, and conceptual representations are, by nature, open to introspection. According to Jackendoff (1987), they are the format in which we think. If the stage of intake follows comprehended input (which is comprised of these conceptual representations), it may imply that intake and any further mental comparisons are also open to introspection. Carroll argues that this scenario might be flawed: the theoretical concept of the black-box LAD does not include conscious introspection. Empirical support has not yet been provided for learners being able to utilize conscious comparison during online processing (Jackendoff, 1987). Upon closer examination, though, Carroll’s contradiction against Gass’ model may simply be an artifact of the researcher’s definition of intake —with Carroll’s model defining intake as a subset of physical stimuli and Gass’s model defining intake as a set of processed structures waiting to be incorporated into IL grammar. Nevertheless, Carroll’s argument raises a point that had not been explicitly pointed out in the previous models: The stage of cognitive comparison (whether it takes place in LAD or at each parsing processor) remains largely automatic and evasive to conscious reflection. At best, the awareness comes only after the fact. This insight is echoed in all the models reviewed.

Thus far, the discussion above has disambiguated the terminological difference of intake and found consensus in the view of comprehensible/incomprehensible input. Also, the unconscious nature of mental comparison is brought to the fore. However, there still remains the issue of noticing and attention. In the models reviewed, noticing and attention have appeared in several places. First, noticing the gap and making cognitive comparisons are viewed by all models as the key to restructuring (Schmidt & Frota, 1986). As mentioned earlier, regardless of the proposed location of such operation, this process has thus far been regarded as automatic and is not open to introspection. As Jackendoff (1987) argued, learners can at best attend and become aware of the resulting representations, but not the actual processes that bring about that representation. The second juncture where noticing and attention have also been discussed is at the preliminary, perceptual stage. Chaudron (1985), Sharwood Smith (1986), and Gass (1997) considered attention as a required element to convert external stimuli into some representations inside of the learners. But Carroll (1999, 2000) asserted that attention should not be used as a blanket term, and an adequate model of input processing needs to take into account the interaction between learners’ current knowledge and attention. In her conception, attention is evoked after the input has been preliminarily processed at the base processor; it is a result of processing, not a prerequisite. Yet within Carroll’s (2000) processors, there also exists internally “a selection function that restricts the number of structures under analysis” (p. 126). It has not been specified by Carroll what mechanisms enable this selection function. Synthesizing all the above views, Chaudron, Sharwood Smith, and Gass’ stance on attention is not actually incompatible with Carroll’s. One possible explanation for this apparent disagreement is that it is an artifact of the way each researcher conceived input processing: Each researcher created his/her model based on a different starting point of processing. More importantly, the diverging views actually highlighted the importance of attention, and it may be so prevalent that it operates before the initial processor, within the processor, and as a result of processing, as suggested by the various models. The importance of attention has already been researched with great interest, as seen in Schmidt’s (1990) Noticing Hypothesis and the substantial body of related studies (e.g., Carr & Curran, 1994; Nissen & Bullemer, 1987; Schmidt, 1990, 2001; Tomlin & Villa, 1991; among others). The way attention...
and input interact has been the central discussion point in so much of the instructional-related research, such as input enhancement, recasts, and processing instruction. From the above four processing models, one can at least conclude that attention (or even awareness) remains a mediating factor in input processing. The appeal of attention in SLA research is that it seems to be more responsive to manipulation and enhancement, whereas the other mechanisms in the acquisitional process (such as the LAD) largely remain beyond conscious control.

**CONCLUSION**

The discussion of the four models above aimed to explore how input is processed and incorporated in SLA. It also demonstrated that the seemingly common terminologies (input vs. intake; comprehensible vs. incomprehensible input) in the models have been conceived by different researchers to encapsulate different components and highlight various aspects of the process as a whole. It has also been found that all four models agree that cognitive/structural comparison is the key to development, regardless of the specific location of operation, though it remains largely beyond conscious control or instructional manipulation. Alternatively, attention may come in as a mediating factor at the perceptual level. There is a substantial body of research available now regarding the actual effect of attention (e.g., Carr & Curran, 1994; Nissen & Bullemer, 1987; Schmidt, 1990, 2001; Tomlin & Villa, 1991). An investigation on the interaction of attention, input, and learners’ knowledge may prove to be fruitful, especially for instructed SLA. In addition, in recent decades, more cognitive and neurological research paradigms have been brought into the field as a different lens to examine the process of SLA (e.g., Baddeley, 2003; Gathercole, Service, Hitch, Adams, & Martin, 1999). The studies on working memory and procedural and declarative knowledge may provide additional evidence and perspectives and may perhaps augment the models reviewed above to produce a more comprehensive picture of the process of SLA.

**REFERENCES**


