CHAPTER-III

THEORIES AND MODELS OF SECOND LANGUAGE ACQUISITION AND THE ROLE OF THE FIRST LANGUAGE

As the title suggests, the purpose of this chapter is to consider the role of L1 in some important prevalent theories of SLA. Three main theories have been dealt with: Contrastive Analysis Hypothesis, Creative Construction and Information Processing Theory. The parameter-setting model is discussed briefly for its role in revitalizing the CA hypothesis through its alignment to Universal Grammar as well as its attempt to converge the Contrastive Analysis and Creative Construction theories. Krashen's Monitor Model, though it assigns a very minor role to the influence of L1 on L2 learning, also figures in the discussion, chiefly because of its prominence in the SLA literature. Other models of learning have been omitted deliberately, either because they do not bear directly upon the matter under discussion or because they are not so well-known.

Contrastive Analysis and the Role of the First Language

The SLA research literature shows considerable difference of opinion regarding the extent of the influence of L1 in second language acquisition. According to Ellis (1986), while some assert the existence of a "constant warfare" between the two language systems (Marton 1981), there are others who advise us to abandon the notion of interference altogether as a natural and inevitable phenomenon in L2 acquisition (Felix 1980b). The notion of
interference in SLA research can be traced to its origin, in the behaviourist learning theory and to its development in terms of Contrastive Analysis Hypothesis. Upto the end of the 1960's, views of language learning were derived from a theory of learning in general. There were hardly any studies based on actual language use by learners and nor were any attempts made to examine the process of SLA through empirical research. The dominant school in psychology was that of behaviourism according to which language was seen as habit formation. And as old habits die hard, therefore, the "grammatical apparatus" programmed into the mind as LI was supposed to interfere with the smooth acquisition of the second. Behaviourist theory predicted that transfer would take place from the first to the second language; it will be negative when two languages share a meaning but express it in different ways; and positive when the first and second language habits are the same, in which case no errors would occur. Thus, differences between the first and second language create difficulty in learning leading to errors, while the similarities between the first and the second language facilitate rapid and easy learning. In the behaviourist learning theory, errors were considered crimes, to be avoided at all cost, and to this end, attempts were made to predict when they occur. It was found that by comparing the learners' native language with the target language, differences could be identified and used to predict areas of potential errors. This procedure was known as Contrastive Analysis. CA was rooted in the practical need
to teach an L2 in the most efficient way possible. It was considered that "the teacher who has made a comparison of the foreign language with the native language of the students will know better what the real problems are and can provide for teaching them (Lado 1957).² Contrastive Analysis had both a psychological and linguistic aspect to it. Whereas the psychological aspect was chiefly based on the behaviourist learning theory, the linguistic aspect was based mainly on structuralist linguistics.

The Psychological Aspect of Contrastive Analysis

The Contrastive Analysis Hypothesis derived from the psychological rationale for exploiting contrastive studies of languages and exists in a strong and a weak form (Wardaugh 1970).³ The strong form of the hypothesis claims that the potential negative transfers from NL on to the TL can be predicted by juxtaposing descriptions of comparable systems and subsystems of the two languages. In other words, it proposed that the main cause or even the sole cause of difficulty and error in foreign language learning is interference from the learner's native language. Contrasts thus identified could be incorporated into pedagogic materials and imparted to FL teachers to minimise the incidence of errors arising from L1 interference. The weak form of the hypothesis claims only to be diagnostic: Contrastive Analysis according to it, can be used to identify those errors which are the result of interference. Hence, the weak version implicitly assumes that every error
is not the result of interference, and allows a less powerful role to the L1 than the strong form of the hypothesis.

**The Linguistic Aspect of Contrastive Analysis**

Although the practice of comparing languages is very ancient, the linguistic fathers of CA are usually recognized as being Fries and Lado. Fries (1945) made the statement that, "...the most effective materials for foreign language teaching are based upon a scientific description of the language to be learned carefully compared with a parallel description of the native language of the learner" (p.9). As already stated before, CA was viewed as a means to predict and thereby circumvent the differences between L1 and L2 features. Various comparative studies were carried out on languages from within the same language family, e.g., Stockwell and Bowen (1965); Stockwell; Bowen, and Martin (1965) using the grammatical model of structuralist linguistics. This emphasized the importance of detailed scientific description of languages based on comparison of different categories that comprise the patterns of a language. The languages were, however, found to vary greatly with regard to their categories. Apart from the disparity of categories, another major problem was that regarding the psychological considerations involved in error production. Stockwell, Bowen and Martin (1965) and Prator (1980a) had proposed that linguistic differences can be arranged in a "hierarchy of difficulty". This claim was, however, based
on the wrong conviction that the degree of linguistic
difference corresponds to the degree of learning difficulty.

Criticism of Contrastive Analysis

By the mid 70's the CA hypothesis came under attack. It
was criticized most menacingly by those, who worked mainly
in a second-language immigrant setting and saw little
unequivocal proof of NL interference, for example, among
Chinese and Spanish children acquiring English in the USA.

Refutation Based on Empirical Data

The existence of non-interference errors in the
language of L2 learners was always recognized except by the
staunchest supporters of the CA hypothesis. Dulay and Burt's
(1972, 1973) research constituted a powerful attack on the
CA hypothesis who, by calculating the frequencies of error
types in the speech data of Spanish-speaking children
learning English, claimed that 85 per cent of the errors
were developmental (errors which do not reflect L1 structure
but are found in L1 acquisition data), 12 per cent unique,
and only 3 per cent interference. On the basis of this and
similar studies, Dulay and Burt argued that children do not
organize an L2 on the basis of transfer or comparison with
their L1 but rely on their ability to construct the L2 as an
independent system, in much the same way as in L1
acquisition. They suggested that interference may be a major
factor only in phonology. However, other research does not
bear out Dulay and Burt's findings as to the exact
proportion of errors that can be put down to transfer.
Percentage of interference errors reported in different studies yielded a mean percentage of 33 percent. The major difficulty in attempts at empirically validating the CA hypothesis has been the lack of well-defined and broadly accepted criteria for establishing which grammatical utterances are the result of language transfer.

Theoretical Criticism

Chomsky's (1959) review of Skinner's "Verbal Behaviour" struck at the psychological basis of theories of language learning. The concepts of stimulus and response, imitation and reinforcement were rejected and ridiculed because they could not account for the creativity of language. It was shown that in L1 acquisition parents rarely corrected formal errors or rewarded correct utterances, and that children were only able to initiate utterances which lay within their existing competence and could not, therefore, learn new habits in this way. Thus, if language could not be explained in terms of habit-formation, then clearly, the central notion of interference was bound to be challenged. This notion, as already mentioned, rested on the assumption that L1 habits intrude into the L2 system. The question arose then, as to what exactly interference consisted of, if it did not involve habit-transfer.

In addition to these criticisms of behaviourist learning theory, there were objections to other aspects of the CA hypothesis. For example, it was criticized for
equating difference with difficulty on the one hand, and difficulty with error on the other. It was pointed out that while difference is a linguistic concept, difficulty a psychological one. Therefore, difficulty cannot be predicted from difference. This assumption that difficulty led to error was also shown to be of doubtful validity. Difficulty and error were shown to be not significantly related. (Jackson and Whitman 1971). Hence, the central claim of the Contrastive Analysis Hypothesis that linguistic difference between L1 and L2 led to error as a result of learning difficulty, was called into question. It was also pointed out that comparison of languages should involve pragmatic as well as structural levels of equivalence. Another problem related with the linguistic aspect of CA was that of accounting for learner variability. Dickerson (1975) posited that CA denies the existence of variability, implying the broad range of L2 learner output which is variable according to the different contexts.

Revival of Contrastive Analysis

A revival of CA however, started to take shape since the beginning of the 1980s, which was the result of the emergence of a new perspective on CA, in keeping with the fresh developments in SLA. The reappraisal sought to answer many of the questions raised by the erstwhile critics of the CA hypothesis.

The Question of Categories
To begin with, the question regarding the comparability of different languages and the dangers of superimposing the descriptive categories of a given language on another, has been answered by James (1980) in the following manner: "First, one does not refer to categories by the same label unless they have at least something in common .... The major defense of the position that languages are in principle comparable, is to insist that comparability does not presuppose absolute identity, but merely a degree of shared similarity" (p.168).

**Small Number of Interference Errors**

As to the very small number of interference errors found empirically in IL production, it was argued that non-occurrence of errors did not necessarily invalidate the prediction but it might imply that the student was avoiding the use of problematic structures (Sridhar 1980). The empirical evidence for avoidance induced by the LI was first provided by Schachter (1974) whose study revealed that Chinese and Japanese learners, whose first languages do not contain English-like relative clauses, avoided using them. The learners' LI, therefore, predicted the extent to which the learners avoided using relative clauses. Bertkau (1974) also found that Japanese students obtained lower marks on the comprehension of relative clauses than Spanish learners. Kleinmann (1978) also indicated that the criticisms of CA hypothesis advanced by Dulay and Burt on the basis of observed error frequency were not foolproof.
Contrastive Analysis is not Necessarily Linked to Behaviourism

CA was associated in its beginning with the behaviourist theory of learning which posited that L1 habits intruded into L2 learning and should, therefore, be necessarily identified and eradicated. However, it has been later incorporated into a mentalistic view of learning. Markham (1985) remarks that transfer from L1 to L2 has a meaning even outside a behaviourist paradigm. Transfer as an aspect of elaboration has always been deemed an essential cognitive process in cognitive psychology literature. Elaboration plays a key role in Anderson's (1983) model leading to effective recall of matter, deductive reasoning and transfer. O'Malley and Chamot (1990) pointed out, "In second language learning, elaboration is of particular interest to researchers studying transfer in bilingual individuals who learn to use the body of prior knowledge originally acquired in the first language to comprehend new information presented in the second language" (p.168).

One of the most prominent cognitive psychologists who strongly emphasizes the importance of transfer is Ausbel and transfer learning is a fundamental component of Ausbel's model (1978). He posits that, "... past experience influences, and has positive effects on new meaningful learning and retention by virtue of its impact on relevant
properties of cognitive structure. If this is true, all meaningful learning necessarily involves transfer. It is impossible to conceive of any instance of such learning that is not affected in some way by existing cognitive structure (p.165)." In view of this, there seems to be no reason to suppose that CA is essentially aligned to the behaviourist theory, and to reject it on the same ground. In fact, transfer theory should be considered an integral factor in the second language learning process.

Universal Grammar and Transfer Phenomena

While Chomsky's theory of Language Acquisition Device (the built-in mechanism in the human mind for language learning), struck at the roots of Contrastive Analysis, paradoxically, it was his own theory of Universal Grammar which was used to account for the validity of CA in later years.

Cook (1985) explains the Chomskyan view of UG in the following words: "The language properties inherent in the human mind make up "Universal Grammar", which consists not of particular rules of a particular language, but a set of general principles that apply to all languages." Thus, UG constrains the form which the grammars of individual languages can take, though this is not done directly by
providing the children with ready-made rules to be incorporated directly into their repertoire. Rather, it sets parameters which must be fixed afterwards according to the input data received by children. These universal parameters delimit the number of options that children require to explore. Children, however, still have to discover which of the various options pertain in the target language. The input data provided by the environment helps them to fix the parameters by selecting the appropriate option.

These rules that the children formulate with the help of the UG are known as the core grammar of their language. However, not all rules are core rules. Every language also contains elements that are not constrained by UG. These comprise the periphery. Related to the concepts of core and periphery is Chomsky's theory of markedness. Core rules are unmarked which means that they conform to the general tendencies of all the languages. Periphery rules are marked, implying that they are exceptional in some way or the other. Marked and unmarked rules, however, are the opposite extremes of a continuum, and it is possible for rules to be less or more marked comparatively.

Markedness theory also provided a basis for solving some of the problems of the CA hypothesis. Specially, it
helped to account for the fact that some differences between the native and the target language lead to learning difficulty, while other differences do not. The basic assumption of markedness theory is that unmarked settings of parameters will occur in interlanguage before marked settings, even if the L2 provides evidence of a marked setting. Thus, it is predicted that no transfer will take place from native to target language when the LI has a marked setting. The most obvious case of transfer is, where the native language shows an unmarked setting and the target language a marked one.

Transfer of L1 Unmarked Forms

Different theories have been postulated for the transfer of L1 unmarked forms. Zobl (1983c, 1984) argues that transfer functions only as an "auxiliary evaluation measure", that is, it takes over when the "projection device" or "triggering" fails to set a particular parameter of grammar.\(^{16}\) In other words, the learner falls back on his LI knowledge only when the L2 rule is obscure, considering the L1 as an "auxiliary evaluation measure" however, assigns only a limited role to transfer in SLA.

Eckman (1977), on the other hand, argued that transfer effects are the most manifest when the L1 setting is
unmarked and L2 setting marked. He postulated his "Markedness Differential Hypothesis" which states that those areas of the target language are difficult to acquire which are both different from and relatively more marked than the L1.

Kellerman (1984) proposed that where the L1 pattern corresponds with a universal developmental stage in SLA, the learner may proceed to that stage faster than learners whose L1s do not have the pattern. He cites Hammarberg (1979), who suggests that English learners of L2 Swedish are likely to miss out the early preverbal negation stage because their L1 does not contain this pattern. Thus, as Kellerman puts it, learners can get a 'leg up' the developmental ladder with the help of their L1. The opposite is also possible. If the L1 contains an unnatural pattern, progress may be slowed down. Such is the case with Spanish learners of L2 English, who can be hindered by pre-verbal negation which is their L1 pattern.

Non-Transfer of L1 Marked Forms

There are arguments and evidence to support the non-transfer of L1 marked forms, although they are by no means definitive. Kellerman (1979) suggests that learners tend to
avoid one-to-one correspondences between L1 and L2 when meanings are considered far from prototypical or universal even though it may result in error. His study showed that learners tend to transfer initially both marked and unmarked features while at more advanced stages they resist transferring marked features. Thus, the acceptance or rejection of L1 peripheral features may be complicated by developmental factors. Stronger evidence for the avoidance of marked features in L1 transfer is provided by Zobl (1984). However, Liceras (1983) presents counter-evidence for the same. There is some evidence from White (1984) for transfer of marked constructions too.

It has been shown that learners are not automata who just transfer NL forms mechanically and unthinkingly. Kellerman (1978) suggested that learners have intuitions about what is transferable from their native language to any specific target language, that is, learners are guided by "psychotypology" which constitutes their belief about the cognateness of the two languages (i.e., NL & TL), and the language-specificity of the linguistic unit in question. His suggestion, thus, is that marked elements of the L1 are not transferable. He defines marked elements as those which are "infrequent, irregular, semantically or structurally opaque, or in any other way exceptional" (p.63).
Parameter-Setting

As mentioned above, the hypothesis formulation arrived at by the child learning his L1 in accordance with the rules of UG grammar is known as parameter-setting. The idea is that language particulars are the reflection of selections from a fixed set of parameters. The great strength of viewing the grammar of a language in terms of universal parameters is that it has been able to associate phenomena which hitherto have been viewed as discrete. Thus, if a language deletes its subject pronouns it will also, among other features allow inversion of its subject NPs. This implies that a long-cherished goal of CA is now at last within reach, i.e., a host of concomitant contrasts can be identified and sets of predictions made immediately, since each such phenomenon implicates others. Hawkins (1986) has shown how it is possible to use such implicational relationships to associate contrasts between German and English syntax that have been known for some time, but hitherto have been considered unconnected. It suggests that a CA can be as powerful as the linguistic theory upon which it rests.

Hence, with the development of transformational syntax there is a scope for parallel development in CA, though Carl James (1990) fears that with the enormous technical sophistication of the modern syntax, "CA is no longer easy to do and is not for the faint-hearted" (p.206).
New Domains of Contrastive Analysis

CA has also been revitalized by widening its perspective beyond the sentence level, so that there exist at present flourishing new domains, i.e., contrastive rhetoric, contrastive pragmatics, and contrastive discourse and text analysis (Hartman, 1980). Contrastive rhetoric of all these has enjoyed the most attention which focusses on contrast between culturally-determined writing conventions. It may be said that it began with Kaplan's (1966) speculations on the general tendencies of argument structure in Semitic, Oriental, Romance, Slav and Anglo-Saxon texts. In fact Kaplan's theory was the starting point for a great deal of serious CA study on Japanese vs. English (Hinds 1983; Jenkins & Hinds, 1987; Oi, 1986; Shimozaki, 1988) and Arabic vs. English (Johnstone-Koch, 1983; Al-Jubouri, 1984).

Contrastive pragmatics involves the identification of cross-cultural differences in speech act realizations, for example, conversational routines such as service encounters, in the pursuit of insights concerning how the native language influences the target language communicative competence (Richards & Sukiwiwat, 1985). In short, the basic idea behind contrastive pragmatics is that languages should be compared in terms of their different linguistic realizations of the same function or alternatively, comparisons should be made of the different functions served by the same linguistic structure in the two languages.
A theoretical framework for the study of NL-induced pragmalinguistic failure was proposed by Thomas (1983). It has focussed on ways of being polite in different language communities. Blum-Kulka (1982) showed that Israelis tend to make their requests more direct than the Americans do. Other speech acts that have been studied contrastively are: apologies (Coulmas 1981), compliments (Wolfson 1989; Herbert 1989), telephone-call openings in France and the USA (Godard 1977), and some others reported in House & Blum-Kulka (1986) and Oleksy (1989).

These studies raised some important questions regarding the extent to which communicative parameters of language are universal or language-specific. Widdowson (1975a) takes a strong stand regarding the universality of specialized communicative functions like those associated with scientific and technical discourse. Elsewhere, he has also pointed out the need to recognize three levels of translation equivalence and to make use of these in the pedagogical context (1974). While under the influence of structural linguistics, equivalence was considered only between the surface features of the two texts, hence giving an incomplete picture of equivalence, Widdowson recommends recourse to semantic equivalence at the deep structure level, as well as pragmatic or rhetorical equivalence between L1 and L2 in order to build upon the previous knowledge of second language learners.
The notable thing about the CA Hypothesis was that though it recognized both positive and negative transfer, most of the focus was on 'interference' or negative transfer of L1 features, resulting in L1 errors. Recently, however, CA has assumed a more positive role with the attention given by scholars such as Ringbom (1987) to the positive and facilitative NL transfer. Ringbom compared the positive as well as the negative transfers made by Swedish Finns learning English and he discovered that the facilitation derived from learning a related language comes in the form of increased comprehension in the early stages. It must be mentioned at the end that notwithstanding all these attempts at revitalizing and updating CA, and adjusting it in the overall SLA theory with a more balanced perspective, a certain suspicion regarding its worth continues to remain.

Mentalist Views of SLA and the Creative Construction Theory

As mentioned above, a mentalist view of language learning asserted itself with Chomsky's review of Skinner's "Verbal Behaviour" in 1959. Chomsky emphasized the active contribution of the child and de-emphasized the role of habit-formation. He attributed the child with an innate language learning capacity which is unique and different from general cognitive mechanisms. The process of acquisition consists of hypothesis-testing by the means of which the child's L1 grammar was formulated on the basis of the universal grammar. The ideas of Chomsky stimulated L1 research activity in the 1960's, which discovered that
children's early utterances were unique in that they were never found in the speech of native-speaking adults. Secondly, development was found to be continuous and incremental and could be characterized by a series of stages.

**Interlanguage and Communication Strategies**

Mentalistic theories and empirical research on L1 acquisition evidently had their impact on SLA. This resulted in viewing the learner's errors as signs of improvement rather than sins to be avoided at all cost. The term interlanguage was coined by Selinker (1972) to describe the L2 learner's language. The linguistic system of the learner was also termed as "approximative systems" (Nemser 1971), "idiosyncratic dialects" or "transitional competence" (Corder 1971), but Selinker's term has gained currency.33

The assumptions underlying the concept of interlanguage were that; (i) interlanguages were distinct from either L1 or L2 at any given point of time, (ii) interlanguage takes shape as an evolving series and, (iii) that interlanguages of different language learners roughly coincide at the same stage of proficiency. The concept of hypothesis-testing was employed as in L1 acquisition to account for L2 learner's progress along the interlanguage continuum. Corder proposed that at least some of the strategies were identical for L1 and L2 learning. He also made the important suggestion that both L1 and L2 learners made errors to test their hypothesis about the language they were learning. Thus, Corder saw the
making of errors as a strategy and a proof of the learner's active internal processing. This notion was obviously in stark contrast to the view of error taken at that time by the CA hypothesis which was then chiefly based on a behaviourist account of learning.

The Role of L1 as a Strategy

The notion of L1 interference was not rejected by the proponents of interlanguage, and they considered it as one of the many contributing factors responsible for SLA. Selinker's (1972) five central processes operating in interlanguage are (i) language transfer (ii) overgeneralization of target language rules (iii) transfer of training (iv) learning strategies (v) communication strategies. Tarone et al (1983) present transfer from the native language as one of the communication strategies, meaning thereby, "the type of negative transfer from the native-language resulting in utterances that are not just inappropriate but actually incorrect by native standards." The process may manifest itself at different levels of language, i.e., phonology, morphology, syntax and lexicon. Corder (1983) mentions "borrowing" from L1 as a "risk taking" strategy. A risk-taking on resource expansion strategy according to Corder is one through which the learners attempt to expand their resources by one means or another in order to realize their communicative goals. Faerch and Kasper (1983) propose "code-switching", "interlingual transfer" and "inter-/intralingual transfer"
as three L1-based strategies. The above L1-based strategies have been drawn from theoretical postulates of early interlanguage literature. Empirical evidence for L1-based strategies is presented in Chapter V.

**Latent Language Structure and Latent Psychological Structure**

As regards its emphasis on hypothesis-testing and internal processes and a continuum of learning, interlanguage theories closely followed acquisition principles. However, the same LAD which accounted for language acquisition in the child could not be held responsible for L1 learning in adults, since this innate learning faculty atrophied at puberty. Hence, the mentalist accounts of SLA had to account for adult L2 learning without the benefit of LAD, and it did this by pointing two different language learning mechanisms: latent language structure, and latent psychological structure (Selinker 1972). Those adults who achieved a native-like proficiency were believed to have continued to use the "acquisition device" which is termed by Lenneberg (1967) as latent language structure, but those who fossilized before achieving such a proficiency fell back on a more general cognitive mechanism, labelled as latent psychological structure. This mechanism was also considered to be genetically determined, but did not involve use of universal grammar and was responsible for the five central processes mentioned by Selinker (1971). The process of language learning involved in latent psychological structures or a general cognitive mechanism which is
responsible for other types of learning too, apart from language, later came to be known as creative construction, and the mechanism responsible for this type of learning was labelled as cognitive organizer (Dulay and Burt 1977).  

Creative Construction Theory and Krashen's Monitor Model

According to the Creative Construction theory, L1 and L2 acquisition are not distinct but follow from the same set of innate principles. In contrast to the traditional CA theory, a CC theory of L2 acquisition claims that prior L1 experience does not influence subsequent L2 acquisition. Rather, the structure of the target language input and the creative construction powers of the L2 learner which all learners share as part of the human competence, are the critical factors in acquisition. Within this framework, the native language serves only a very peripheral role, "integrated into some of the organizing strategies used by a learner to acquire an L2 (Dulay & Burt 1977, p.159).

Dulay and Burt's Creative Construction theory and Krashen's Monitor Model (1981a, 1982) are closely related to each other. Though Krashen does not assign an important role to the learner's L1 in SLA, it merits some attention because of its considerable prominence in the SLA literature, inspite of its being extremely controversial. The Monitor Model rests on the following five central hypotheses.
i) There is distinction between acquisition and learning. Acquisition takes place subconsciously in a natural setting when the focus is on meaning, while learning is derived from conscious study with focus on form. Both comprehension and production are based on the "acquired" reservoir of knowledge while learnt knowledge only serves as a Monitor.

ii) The learners follow a fixed and predictable order of acquisition of grammatical structures.

iii) Learnt knowledge can only serve as a Monitor to edit the learner's performance. There are three conditions for the use of the Monitor: sufficient time, focus on form, and knowledge of the rule concerned. Editing can take place also through acquired knowledge, but this monitoring is different, and is referred to with a small "m".

iv) For acquisition to take place, the presence of "comprehensible input" is necessary which is just a step beyond the learner's current level of competence. This formula is denoted as "(i+1)".

v) Finally, Krashen speaks of "affective filter" as a causative factor in SLA which controls the learner's intake of input. Positive emotions like motivation and self-confidence constitute a low affective filter, allowing high intake; whereas negative emotions like lack of confidence, anxiety and low motivation comprise high affective filter, resulting in low intake. The first language plays a very minor role in Krashen's model. Krashen rejects the notion of
LI interference and views the role of LI as a performance strategy. The learners make recourse to their LI when they lack an item in L2. The strategy involves planning their utterance employing LI, and then replacing it with L2 lexical items. In this process they also employ the Monitor to make little corrections.

Interface, Non-Interface and the Role of LI

As mentioned above, Krashen argues that acquired and learnt knowledge reservoirs are distinctly separate from each other and acquired items can not be transformed into learnt ones or vice-versa. This stance is known as the "non-interface" position in SLA literature. (Ellis 1986). The opposite of it, that is, the interface position, maintains that there is flow from one knowledge type to the other. A weak interface position has been proposed by Seliger (1979) who considers internalization of rules as different from formal rule learning, but believes that pedagogical rules help in internalization when the context is proper and also, that rules may facilitate the acquisition of those features which though acquired, are still shallow.37

Contrary to this, Stevick (1980) takes a strong interface position, believing that transfer of knowledge takes place from learning to acquisition and vice-versa. There are other views too, which support a strong interface position (e.g., Bialystok & Frohlich 1977; Bialystok 1979; and Bialystok 1981) and suggest that explicit knowledge can be turned into implicit knowledge in two ways: first through
"unconscious acquisition" and secondly through the automatizing of explicit knowledge through practice. McLaughlin also believes Krashen's acquisition/learning distinction to be "ephemeral" and substitutes it with Schneider and Shiffrin's (1977) distinction between "controlled" and "automatic" processing. Controlled processing requires active attention, while automatic processing takes place without active control or attention. Significantly, controlled processes themselves become automatic after extensive practice, and there is no necessity of postulating two detached knowledge types. Sharwood-Smith (1981) builds on the work of Bialystok and McLaughlin and develops a full interface model to account for the role of formal instruction in SLA which helps in consciousness-raising of metalinguistic knowledge. According to his model the learner can produce L2 utterances through three different sources: implicit knowledge, explicit knowledge or a combination of these two. Explicit knowledge can become completely automatized through repeated performance. The weak interface position can account for both the failure to find any positive effect on the route of SLA and for the evidence that classroom instruction does affect the success of SLA. However, the strong interface position can only explain the evidence for success but is less comfortable regarding the effect on the route of SLA. Distinct from these interface and non-interface positions a third position is taken by Ellis (1986): the "variability" position which maintains that the kind of language use that
the learners engage in, determines the kind of knowledge that they acquire.

The interface and non-interface argument is important in a discussion regarding the role of L1 in formal instruction. According to Krashen's non-interface position, focus on the formal elements of language does not lead to acquisition and metalinguistic awareness (knowledge about the rules of language) only serves as a Monitor. Since, recourse to learners' L1 in the classroom, specifically translation from L1 to L2 or vice-versa, is often employed to compare and contrast the formal features of the two languages, it would be considered of little value according to Krashen. However, it has a significance in the models provided by Bialystok, and McLaughlin; and Sharwood-Smith's concept of consciousness-raising is specially relevant to the idea of using L1 as a source of bringing out the similarities and differences of the two language systems.

**Convergence of Contrastive Analysis and Creative Construction within a Parameter-Setting Model**

Empirical facts support some evidence of creativity and some evidence of L1 interference. The Creative Construction theory is supported by findings made by Dulay and Burt (1974, 1977); Dulay, Burt and Krashen (1982); Cook (1973); d'Anglejan and Tucker (1975); Hyltenstam (1977); Mazurkewich (1985); and Gass (1980) among others. On the other hand, the CA hypothesis is corroborated by findings by Selinker (1969); Taylor (1975); Rutherford (1983); White (1985);
Duskova (1969); and Liceras (1983) among several others. It is clear that CA and CC are in fact not mutually exclusive as considered traditionally.

According to Flynn, contrast and construction, both significant in L2 acquisition, should be accounted for in a sound SLA theory, both psychologically and linguistically. Flynn, along with several other researchers (e.g., Flynn (1984, 1987a,b,c); Liceras (1983, 1985); Haegman (1985); and White (1985) have recently proposed that a theory of Universal Grammar, specifically a parameter-setting formulation provides the necessary theoretical framework which allows a reconciliation of the traditionally disparate components of contrast and construction.

As suggested by Chomsky (1981), a theory of UG specifies that abstract and linguistically significant principles underlie all natural languages, which also define the initial prelinguistic state of the child's mind and "restrict the class of attainable grammars and narrowly constrain their form." These principles of UG, thus determine the basic grammar of the language acquired by the child, and comprise the essential faculty with which all individuals are in general uniformly and equally endowed. Chomsky also specified that a number of these principles are associated with parameters. Experience with the input data helps in the setting of the parameter. Thus, while setting of the parameter in a particular way yields one language, setting it in a different way yields another.
However, UG and its properties characterize L1 acquisition and do not directly predict the process of L2 learning. But, if principles of UG do provide for a biologically determined language faculty which is sufficient to explain how language acquisition is possible, then UG should also be able to account for L2 acquisition in some way, under the assumption that language faculty does not change substantially over time.

Within the particular model developed by Flynn (1984, 1987a), it is argued that the essential faculty for language proposed in L1 acquisition is also critically involved in L2 acquisition. The role of L1 experience is explained in the following way within the model:

A parameter may be defined as a general organizing principle for grammars of all languages. The particular value of a parameter will vary from one language to another, but the underlying principle remains the same across all languages. If the principles responsible for L1 acquisition also determine and account for L2 acquisition, the values for these parameters will match with those of the L2 parameters. In cases where there is disparity between L1 and L2 values, new parametric values will have to be assigned for these principles. In cases where the values match, no such assignment will be required, and acquisition will be facilitated as a result. On the other hand it will be hindered where there is mismatch of these values.
Flynn also proposes that L2 learners first establish the "basic structural configuration" of the L1 before they can use this configuration to guide them in learning the features of L2 grammar (Flynn, 1987a, p. 30). In the case of a match of values, since the structural configuration will have already been established it will manifest itself at an earlier stage than in the mismatch case.

**Creative Construction within Flynn's Model**

This theory is consistent with a CC theory of L2 acquisition in that both L1 and L2 are hypothesized to be constrained by a similar set of principles. However, it differs from a CC theory in that the L1 experience is claimed to affect L2 acquisition more profoundly than is normally proposed by a CC theory: That is, a CC theory claims that L1 experience emerges in L2 acquisition only in the form of some type of vague astructural organizing principle within Flynn's theoretical framework, however, L1 experience emerges significantly in L2 acquisition in terms of its role in determining whether new values must be assigned to parameters to cohere with the L2 values.

**Contrastive Component within Flynn's Model**

Flynn's model is also consistent with a CA theory in that the L1 experience is important. However, it differs from the traditional CA theory in various ways. Firstly, learning within the proposed theory is not claimed to be a habit-formation process. Rather, consistent with a general
theory of UG, learning is claimed to be guided by innate principles of language recognition. The interference data provides evidence that the L2 learner is trying to deduce the value of certain parameters of language organization and is attempting to organize the L2 grammar in accord with these principles.

In brief, Flynn argues that the parameter-setting model of Universal Grammar proposed by her for L2 acquisition provides the scaffolding necessary for an integration of these two components within one explanatory account. Consistent with CC, L2 learners within this model use principles of UG, established in L1 acquisition in the construction of L2 grammar. However, when values of parameters associated with these principles differ between the L1 and the L2, learners assign a new value to cohere with the values for the new target language. The L1 experience counts in determining whether such a new assignment of a parametric value is necessary. This aspect of the model is consistent with a traditional CA theory of L2 learning. Flynn supports the model by empirical evidence derived from a study which investigated the role of the head-initial/head-final parameter (Stowell 1981) in adult L2 acquisition of pronoun anaphora. Two groups of adults - L1 speakers of Spanish, a head-initial language, and L1 speakers of Japanese, a head-final language, were studied in their elicited production of English, a head-initial language. Results indicated that both groups of learners used the head-initial/head-final parameter as a source of
structural organization for the L2. Results also indicated that Japanese speakers are sensitive from early stages of acquisition, to the mismatch in head-direction in English and Japanese, and that they assign new values to this parameter to cohere with the target L2 value.

Flynn's parameter-setting model is one of the attempts to relate the common cores of the two apparently contradictory theories of learning. The picture of SLA which is emerging through current research does suggest that the process of creative construction based on new input as well as reliance on prior knowledge including the knowledge of L1 are both two important aspects of L2 learning. It is expected that further research in future will confirm these tentative propositions before long.

**The Information Processing Model of SLA**

The information processing model of SLA has been derived from cognitive psychology and is based to some extent on the information processing of artificial intelligence. The brief account of the model presented here, as well as the role of L1 transfer in it, are based chiefly on O'Malley and Chamot's (1990) description of Anderson's model (1983, 1985) of learning, and its application to SLA which is considered as similar to the process of acquiring "complex cognitive skills" described by Anderson. It should be noted that in Anderson's model, learning a language is not different from other types of learning.
According to the information processing model, knowledge or information is stored in memory in separate reservoirs: short-term memory, (STM) and long-term memory (LTM). Any new information reaches the STM first, but since the capacity of STM is very limited, it can handle a very small number of items at a time and that too, for a very brief period. Long-term memory on the other hand constitutes a sustained store of information, to which information processed in STM initially, is transferred for the subsequent stage.

Furthermore, knowledge is represented in the memory in two forms: declarative knowledge and procedural knowledge. Declarative knowledge is static, e.g., names, facts, definitions of things and rules. Declarative knowledge is represented in memory in terms of meaning rather than exactly replicated external events and the most significant mode of storing information through propositional representations. The basic unit or element of the propositional network is a node which can also be called an idea. The links in the network of nodes are similar to associations between these ideas. A larger unit of meaning than the proposition is called a schema. A schema is a configuration of interrelated features making up a concept. Schemata are useful because they help in making inferences, and organizing and understanding new information on the basis of the associative networking within them. Thus, an item of new input tends to be associated with the already
present schematic associations in the mind and acquires a broader and deeper meaning on the basis of prior knowledge.

**Declarative Knowledge**

Declarative knowledge is divided into two types: world knowledge and linguistic knowledge. Two special types of schemata represented in world knowledge are scripts and story grammars. While scripts are situation-specific procedures or particular event sequences, story grammars represent discourse organization of narratives, stories and fables. Linguistic knowledge which is also represented in the form of schemata or propositions, consists of a lexicon of word-meaning and a body of syntactical rules.

These two types of knowledge representations, i.e., world knowledge and linguistic knowledge aid top-down and bottom-up processes respectively. The terms top-down and bottom-up are used in different contexts in SLA literature. In the context of reading, while top-down processing entails use of past experience and knowledge store in order to interact with a text, bottom-up processing implies analysis of each individual word and syntactical structure.

**Procedural Knowledge**

Whereas declarative knowledge is static, procedural knowledge is dynamic and refers to the "how to" aspect of information in memory. The ability to understand and produce language on the basis of known facts and rules is a type of procedural knowledge. While declarative knowledge can be
acquired quickly, procedural knowledge is acquired gradually and only with extensive opportunities for practice. Procedural knowledge has been given a central role in Anderson's model of the learning of "complex cognitive skills" and has been incorporated in what Anderson has called "production systems". Anderson has pointed a unitary theory of the mind or common cognitive system for all higher level learning. His position, however, is directly opposite to Chomsky's (1959), who proposed unique mental faculties inherent in the human mind for language acquisition.

In its most basic form, a production has a "condition" and an "action". In other words it can be represented as IF-THEN conditional sentences where for every IF clause, there are several THEN clauses to choose from. Anderson (1980) gives the following example of a production for pluralization: "IF the goal is to generate the plural of a noun, and the noun ends in a hard consonant, THEN generate the noun +S". It is important to note that condition-action pairs such as this are initially represented in declarative form and are converted into procedural form or production only after extensive practice.

**Transition from Declarative to Procedural Knowledge**

This leads one to the question as to how one reaches the automatic proceduralized stage from the rule-bound declarative knowledge in the performance of complex cognitive skill. For this transition to take place, Anderson
has proposed three stages of skill acquisition: the cognitive, the associative and the autonomous stage. At the cognitive stage, knowledge is in a conscious and declarative form, and it can be described verbally by the learner, but this knowledge by itself is inadequate for the performance of the skill and performance when taken up, tends to be laden with errors.

During the associative stage of learning, errors are detected and eliminated and associations between various elements of the skill are strengthened. Basically, during this stage, declarative knowledge is converted into procedural form, but the declarative representations initially formed, do not always last.

At the third stage, the execution of the skill becomes virtually automatic, and errors inhibiting successful performance disappear. There is much less demand on short-term or working memory, or consciousness at this stage. It is important to note that skilled performance can be improved only gradually. It implies that while declarative facts can even be learnt in one trial, a complex skill like language acquisition requires a long period of practice.

The cognitive and autonomous stages of Anderson are congruent with Schneider and Shiffrin's controlled and automatic processing (Schneider and Shiffrin, 1977). Cognitive tasks may be conceived of as involving controlled processes, requiring the attention of the learner, and autonomous tasks involve automatic processes, not requiring
the attention of the learner. According to Schneider and Shiffrin, during the performance of the controlled processes, important portions of the STM are consumed by the execution of the skill, and there is little capability remaining for the performance of additional skills. In automatic processing on the other hand, nodes in the LTM are activated and action sequences or "production" can be performed without demands on the STM. The activation of nodes in the LTM in automatic processing is a learned response, and is built up over an appreciable period of time.

Transfer of L1 Knowledge and Skills in an Information Processing Paradigm

Transfer of Declarative Knowledge

Cognitive theorists have debated whether bilingual individuals have two separate stores of information in LTM, for L1 and L2 or a single store of information accompanied by selection mechanisms to serve both L1 and L2 (McLaughlin, 1984). If individuals have separate information stores, then appropriate information would be selected by language users from either store and transfer of information from L1 to L2 would be difficult in this case, because of the segregation of L1 and L2 memory compartments. This argument for segregated propositional networks and schemata associated with each language is consistent with domain-specific language acquisition according to O'Malley and Chamot (1990). Learners may acquire one or more domain-specific language uses (e.g., language consistent with
shopping encounters or family interactions on different social groups), but be ineffective in communicating in other domains because of the highly specific nature of the language involved. Anderson, however, postulates a unitary information store for both L1 and L2. In his theory, information stored in memory has a meaning-based representation independent of any specific language, and is stored as declarative knowledge through either propositional networks or schemata. Anderson's notion of common propositional networks and schemata for both languages is consistent with Cummins's (1984) common underlying knowledge for cognitive and academic language proficiency (CALP) for bilinguals. Cummins proposes that L2 learners need not relearn what they have originally learnt through L1, but their L1 knowledge can be transferred and expressed through the medium of L2. Cummins's view is consistent with Anderson's suggestion that nodes are based on meanings instead of direct replication of language. Nodes that represent meaning in LTM are non-language specific but have built-in features that select one language or the other.

Anderson's theory also raises the question whether there is variation in the ease and effectiveness of transfer from L1 to L2 declarative knowledge. As already mentioned earlier, declarative knowledge is represented in memory as schemata or organizational frameworks. Two major types of schema are, organization by natural categories, and organization by events. Natural category schemata are based
on real world phenomena, such as classification of plants, animals, minerals and other aspects of the physical world. A natural category schema would appear to be easily transferable to L2 linguistic form. Schemata organized, by events include both personal recollections of event sequences and the sequence of events in a story. Since these story scripts are strongly influenced by culture, an L2 learner might face difficulty in transferring these L1 story scripts because of differing cultural expectations. Another type of schema referred to by Anderson as social cognition might also be difficult to transfer because persons organize their knowledge about individuals or groups according to certain perceived characteristics, which in turn are based on their respective cultures.

The educational implications of the above for L2 learning are that concepts related to natural categories such as science, mathematics and technical subjects may be easier to transfer to the L2 than concepts related to the L2 culture.

Transfer of Metalinguistic Knowledge

Metalinguistic knowledge of L1 can also be transferred to L2 like other forms of declarative knowledge. One particular way to describe metalinguistic awareness in cognitive theory is as a new schema constructed to link independent schemata in the L1 and L2 that refer to the same domains. In other words, an individual with domain-specific implicit knowledge in two languages may begin to see the
different ways in which the same concepts can be expressed in the L1 and L2. The learner then establishes a new schema that differentiates applications of each language to identical concepts. Another proposition of transfer of metalinguistic knowledge which is in keeping with Cummins's (1984) notion of common cognitive and academic language proficiency is like this: because meaning is not inextricably linked to lexical or syntactical aspects of a particular language, underlying meanings may have more than one language schema linked to them. For example, a person could have a single propositional representation for making requests, but will link it to different language schemata for L1 or L2, based on an analysis of the different ways in which requests are made in the two languages.

Cummins, however, argues that the cognitive benefits of transfer in bilingualism can be had only when a learner has reached a "minimum level of bilingualism" which is known as his "threshold hypothesis". He also argues that the level of competence achieved in L2 is dependent on the child's L1 proficiency at the beginning of intensive L2 exposure. This is known as Cummins's "interdependence hypothesis". Failure to meet Cummins's two conditions could undermine the ability to transfer information from L1 to L2, in that schemata established in L1 may not be equipped with sufficient internal associations, to provide a base for L2 acquisition through transfers.

Transfer of Procedural Knowledge

197
The basic unit of procedural knowledge or skill is called a "production" in Anderson's (1983, 1985) model. As mentioned previously, a production has a "condition" and an "action", i.e., it is goal-oriented, and for every goal to achieve, there are several conditions to choose from. Different sorts of rules, i.e., sociolinguistic, discourse, strategic and grammatical competence have to be applied in order to materialize a communicative goal. These are presented in the memory initially as declarative knowledge. Thus, L1 procedural knowledge which is in declarative form in the beginning, is transferred from L1 to L2 as other types of declarative knowledge. The transfer of procedural knowledge, however, has not been elaborated fully by O'Malley and Chamot. They speak of difficulty arising in the transfer of procedural knowledge because of the load added to short-term memory processing requirements which already are burdened by trying to decode new language.

In the present researcher's view, more important however, is the lack in Anderson's model of a unitary system of representation of procedural knowledge which can be available to both L1 and L2 skills, as is the case with declarative knowledge, represented in memory in terms of non-language specific meaning-based propositions. It should be noted though, that if such a common representational system is achieved for procedural knowledge too, Anderson's model will end up being something like Universal Grammar. O'Malley and Chamot (1990 p.84) admit that "... with regard
to procedural knowledge, the theory raises questions about the basis of production systems and language based rule-systems and the commonality of production systems in the L1 and L2" (p.84).

To sum up, where Anderson's model succeeds in explaining many SLA constructs, it fails at the same time to explain many other. Not only this, it raises important questions, first by considering all language learning a conscious process, unless it is automatized, and secondly by considering all SLA the result of rule-formulation. Both these propositions can arouse strong objections, but since these considerations are out of the scope of this work, they will not be discussed here.

**Conclusion**

In conclusion, the chapter has dealt with three theories of SLA: Contrastive Analysis, Creative Construction and information-processing with respect to the role they assign to L1 in learning a second language. The first two of these are the most widely recognized theories of language learning. CA, which assigns an important role to L1 in SLA was first grounded in the behaviourist theory of learning, has made a comeback in a mentalist framework in recent years. Some researchers have even tried to converge it with the Creative Construction theory. The Creative Construction theory is the most widely accepted theory of SLA at present, though its expression has taken many different forms. Krashen's Monitor Model is one of the best known models of
SLA endorsing this theory, but it does not allow a significant role to L1 in SLA. This does not undermine the significance of the mother tongue, however, since Krashen's model has itself been criticized severely on several grounds. The information-processing theory, which has attracted considerable attention for some time now, has been discussed in the end in a little greater detail because of its being new to SLA literature. As described above, it attributes an important role to transfer of L1 knowledge and skills to L2.

Apart from the above three, there are various other theories and models of SLA such as Schumann's Acculturation Model (1978); Giles's Accommodation Theory (1982); the Discourse Theory proposed by Hatch (1979); Ellis's Variable Competence model (1984a) and Lamendella's (1978b) Neurofunctional theory of SLA. However, they have not been discussed here because of either of the two reasons: first, they were not found to be directly relevant to the work; second, they have not acquired the same degree of prominence in SLA as the models and theories discussed.
REFERENCES


H. Dulay and M. Burt, "Should We Teach Children Syntax?" Language Learning, XXIII (1973), 245-58.


10. C. James, Contrastive Analysis (Longman, 1980).


Krashen (1982), op.cit.


