	Applications of Universal Grammar (UG) in the EFL/ESL Classroom
	Kirkwold, Lorne O
	, 32: 59-94
	2005-11-30

Applications of Universal Grammar (UG) in the EFL/ESL Classroom[1]

Lorne O. Kirkwold

TABLE OF CONTENTS

Abstract

The article proposes Stern's (1983) framework for classifying issues related to instruction in order to ascertain the relevance of Universal Grammar (UG) in the EFL/ESL classroom. Issues related to: (a) L_1 transfer; (b) teaching rules and giving error correction versus presenting structures by analogy; and (c) the extent to which the focus should be content rather than language are discussed in this article, particularly as UG pertains to them. The article will be of interest especially to foreign language teachers and also second language acquisition (SLA) researchers. The author draws on some examples in English and French, but then presents his conclusion along with further issues he raises based on his recent experience in universities of the Sapporo region.

Keywords: *ESL, EFL, Universal Grammar*

INTRODUCTION

H. H. Stern was a highly respected language educator at the University of Toronto's reputable Ontario Institute for Studies in Education (OISE). In his 1983 *Fundamental Concepts of Language Teaching*, he identified what he perceived as three central issues of language learning: (1) "The L_1-L_2 connection;" (2) "The explicit-implicit option;" and (3) "The code-communication dilemma" (pp. 400-405). These issues as he enumerated them may prove useful as a checklist for classroom teachers for evaluating theoretical framework. In particular, the purpose of this paper is to consider applications of Universal Grammar (UG) which are immediately germane in EFL/ESL instruction. As a definition of UG, Pinker (1995) includes this entry in his glossary:

> The basic design underlying the grammars of all human languages; also refers to the circuitry in children's brains that allows them to learn the grammar of their parents' language. (p. 483)

The concept may require further explanation at this point. Readers are referred to Chomsky (1998) regarding the nature of the "mental organ" (p. 180), to Archibald and Libben (1995) for L_2 issues related to UG, and to Larsen-Freeman and Long (1991) for an overview of L_2 research.

As for the intent of English as a foreign language (EFL), this is the usual designation for English taught where the language has restricted use in the community outside the classroom. This would be the designation for most instruction in Japan. English as a second language (ESL) is the usual designation for English taught to non-native speakers in countries where the language has official status: the United

States, the United Kingdom, Canada, Australia, and New Zealand. These terms are relative. Take the example of the Province of Québec, where the French-language Charter is intended to protect the Québécois identity essentially through exclusive unilingualism. At the same time, the Federal Government of Canada's official policy is to promote bilingualism. I have chosen the example of Québec to illustrate the extent to which EFL and ESL are relative terms. Given the immediate access these days of the World Wide Web, virtually all languages, irrespective of status, can be accessed around the world. This technological development may be a considerable influence in the trend towards globalization, thereby resulting in more thoughtful definitions and descriptions about where English is appropriately taught as ESL or EFL. Throughout this article, the first language, or mother tongue, will be referred to as L_1. Similarly, the term second language includes foreign languages, and will be designated L_2.

1. THE L_1-L_2 CONNECTION

Stern encapsulates two debates with this heading. First, in what circumstances should a language be learned solely through the L_2 target language? Second, how does the presence of the first language influence the development of the second?

1.1 L_1 as a Medium of Instruction

At the first level, the desirability of exploiting the first language as a means of instruction has been debated at length. Traditionally, the two methods representing the two sides of this debate have been grammar-translation and the direct method. There are also recent methods which reflect the polarity of this debate. Counselling-learning and Suggestopedia, for example, include procedures which rely heavily

on the mother tongue[2]. On the other hand, procedures described in the Krashen and Terrell's (1998) Natural Approach and Asher's (2000) Total Physical Response reflect the underlying assumption that the second language by itself is the best way to effect acquisition. As for the discussion of this level of the debate about the L_1-L_2 connection, bilingual teachers may find UG useful in terms of the predictions it makes, particularly in terms of corrective input, when structures differ between L_1 and L_2.

1.2 The Influence of L_1 on Emerging Interlanguage

1.2.1 Chronological overview

As for the second level of the L_1-L_2 connection, this part of the debate has focussed on the influence of the first language in second language development. The chronology which follows will show the extent to which issues related to L_1 on L_2 development have been pursued and then left unresolved. Reviewing the literature published from the middle of the fifties until the end of the seventies, we can take account of the change in course with regard to the explanation of the mother-tongue influence on L_2.

During the 60's, the contrastive analysis hypothesis was the linguistic approach endorsed by many American structuralists. Learner errors were explained as negative transfer from the mother tongue. Lado (1957) is popularly cited for this quotation:

> ... and since the learner tends to transfer the habits of his native language structure to the foreign language, we have here the major source of difficulty or ease in learning the structure of a foreign

---

[2] For descriptions of these methods, please refer to Richard and Rodgers (1986).

language. Those structures that are similar will be easy to learn because they will be transferred and may function satisfactorily in the foreign language. Those structures that are different will be difficult because when transferred they will not function satisfactorily in the foreign language and will therefore have to be changed. (p. 59)

Contrasting L_1 and L_2 structures would supposedly determine those of greatest difficulty to master. Consequently, according to structuralists, learning a new language would require building up target structures to attain accuracy.

At the beginning of the 70's, second language acquisition researchers adopted methods developed by L_1 child-language researchers: Brown (1973) and deVilliers and deVilliers (1973). An example of such was a study conducted by Dulay and Burt (1974), who developed the bilingual syntax measure (BSM). Their instrument was presumed to rank accuracy on the morphemes of English (-ing, s, ed, 's, etc.). Their work, along with that of Bailey, Madden and Krashen (1974), made a case that, similar to native-speaking children, L_2 learners essentially all acquire English in the same developmental order, albeit a different one from L_1 child language. Accuracy on certain morphemes of the language was expected before others. Based on this research, Dulay, Burt, and Krashen (1982) present their case for "creative construction" in *Language Two*.

By 1975, L_2 acquisition was suspected not to be the neatly ranked sequence generated by the BSM. Larsen-Freeman (1975) administered the BSM with four other tasks to her subjects. As expected, the BSM ranked morphemes in the predictable way. However, only one of the other tasks ("imitating") ranked an order of difficulty correlating

significantly with the BSM order. Furthermore, when a single language group experienced greater difficulty than another with a given morpheme, contrastive analysis could often explain the error.

After Larsen-Freeman's study, some inherent flaw in the BSM procedure was suspected. In 1976, Hakuta published further counter evidence, among other extensive analyses, against the BSM morpheme sequence. In a 60-week longitudinal study of a 5-year-old Japanese learning English, Hakuta applied Brown's morpheme acquisition criteria, and found that his subject acquired morphemes in an order different from the BSM sequence. He also reported that Gillis, studying two Japanese subjects, obtained results differing from previous BSM studies.

By the late 70's, literature began to reflect the limitations of studies ranking the acquisition of morphemes. Andersen (1977, 1979) wrote critically about the limitations of a morpheme-accuracy study he had conducted preferring instead to focus on the following three English structures: NsN (possessive morpheme-'s), NN and N of N, which develop in the target language in the place of N de N from Spanish. Interpreting his results gathered from 89 Hispanic L_1 subjects, he observed that certain L_2 errors are neither strictly interlingual nor intralingual. They can only be explained by the application of an L_1 rule and L_2 rule simultaneously. For example, Andersen claimed the error "wife-house" (for "house-wife") is the result of an English NN construction with interference from Spanish word order.

Huebner (1979), also writing critically about his tabulation of accuracy on articles, endorsed instead his longitudinal study of an uninstructed adult Laos refugee learning English to analyze the contexts in which the subject produced articles--morphemes which remained unacquired throughout the one-year study according to

Brown's criteria.   Huebner wanted to explain the rules of the learner's interlanguage grammar for this morpheme.   He constructed a table of distribution based on the occurence of articles in spontaneous conversation at four times in a year.   From his analysis, Huebner concluded that a progression could be attributed to changing rules governing the subject's interlanguage.   Huebner speculated his subject's production was first governed by topic prominence (a feature of L_1, Hmong) and later showed emergence of subject prominence (a feature of English).

By the end of the 70's, much criticism had been levelled against the study of morpheme accuracy as an indicator of common development of English for L_2 learners.   This was because the analysis required consideration of L_1 influences at the same time.   It was commonly thought that although morpheme-accuracy studies may have tentatively been indicative of a developmental order, by themselves, they could not explain it.

1.2.2   UG Framework for Explaining L_2 Development

Through this chronological overview, we can better understand what we would expect of a comprehensive theory which could explain the underlying mechanism and predict the developmental order of a second language.   What we can glean from this chronology is that a comprehensive theory would ideally explain the progression of interim grammars as well as the order common to all learners and, in the case of L_2, the influence of the mother tongue.

1.2.2.1   UG Framework for Explaining a Common Developmental Order

With respect to an order common to all learners, there would be universals in UG theories which remain constant for both L_1 and L_2.

At the same time, since language competence implies more than knowing rules, Roeper (1981) outlines a theory of "triggers," which effect development. In her discussion, White (1987) includes four kinds: (a) cognitive, (b) maturational, (c) semantic, and (d) linguistic. Let us consider an example of each trigger. This may provide insight into a developmental path which would be common to all learners.

(a) Cognitive.

Consider a typical elementary classroom activity in which the teacher invites the anglophone pupils to volunteer examples of concrete nouns and abstract nouns. I suspect that when children are able to give examples such as "book" for for the former, and "fright" for the latter, they have also acquired enough grammar implicitly to make judgments about what is allowed in spoken language. Consider the following contrast:

(1a) My friend gave me a book.   (2a) My friend gave me a fright.

(1b) My friend gave me two books.   (2b)*My friend gave me two frights.

I have used 'book' as an obvious example of a count noun in (1a) and (1b). In (2a), 'fright,' appears to be a count noun since it occurs with the indefinite article. In (2b), however, 'fright' resists pluralization and is therefore marked with an asterisk for ungrammaticality[3]. At the same time, a pupil in the classroom activity described earlier presented with the examples of this contrast could identify the three allowable sentences and the unallowable and therefore ungrammatical one. In this hypothetical example of maturity and cognition, I suspect the child could produce the right determiner (a, an, the, some, any) reasonably

---

[3] I assume (2a) is best explained as a collocation.

accurately before a singular count noun and differentiate before a mass noun, in which position a null determiner may be required according to context. In place of grammatically tests and questioning the child about rules of mass and count nouns, I suspect the same end may be reached through the quicker identification of concrete and abstract nouns. From this simple test (both linguistic and cognitive), the teacher could also make inferences about native-speaking students' ability to use the standard adult article system in front of mass nouns and count nouns.

We could also expect various levels of cognition to influence second language development. For example, certain levels of cognitive maturity will permit the L_2 learner to progress through expected stages of development, say hypothetically from count noun to mass noun to binary noun. At the same time, the lack of awareness about "things" (for count nouns) and "stuff" (for mass nouns) may result in certain regular peculiarities in the interlanguage. An adult learner will have sufficient cognitive maturity to deal with this categorization analytically. It is possible that with the benefit of explanation, this learner could eradicate these deviant forms. In the same way, a later step may be "binary nouns" (Greenbaum & Quirk, 1990, p. 98). Counting nouns such as shoes, pants, and glasses requires the lexicalization "two pairs of ..., three pairs of ...." Negative input may therefore prove useful. Beyond this, there are infrequent forms of lexicalization such as "a stalk of celery" and "a head of lettuce."

On the other hand, children whose mother tongue does not differentiate between mass nouns and count nouns in the same way as English may be expected to produce errors until they, like adult L_2 learners, have reached a level of cognition which permits analytical metalinguistic awareness which allows distinguishing between mass nouns, count

nouns, and binary nouns. Until such time, negative input would likely not affect the interlanguage.

(b) Maturational.

In an attempt to reconcile the logical problem of language acquisition with the data observed in the L_1 child language studies, Felix (1984) has posited a maturational schedule. Ideally, he posits a UG circuit which consists of principles $P_{(1)}$, $P_{(2)}$, $P_{(3)}$, ... $P_{(n)}$. As the child progresses across a time line of $T_{(1)}$, $T_{(2)}$, $T_{(3)}$ ... $T_{(n)}$, the respective principles become accessible and initiate the development in the child's interim grammar. In other words, a child at $T_{(3)}$ will have access to principles $P_{(3)}$, $P_{(2)}$, and $P_{(1)}$. As this child progresses along the time line, the ever increasing number of principles influencing this individual's interim grammar will make it increasingly adultlike.

This model has interesting implications for second language development. If the developmental sequence is essentially linear, there are two possible scenarios for adult L_2 learners. One possibility is that they will begin the circuit again at time-zero and wait for the emerging principles to effect interlanguage development according to the set agenda. On the other hand, perhaps maturational development through L_1 would have already ensured that the basic principles governing language are fully accessible to the adult learner. In this case, we could expect parametric setting in tandem with the basic principles to explain transfer from the first language.

(c) Semantic.

A semantic theory of language is implicit in the communicative language movement of recent years. This development keeps the focus on the importance of meaning in language. The incorporation of

a lexical component would therefore complement the more elaborate grammatical side of UG.   The dissemination of semantic principles has come from outside UG.   However, the need for the consideration of semantics is quickly apparent when it comes to the most cursive examination of dative movement (also called "dative alternation"). Presented below are contrasts similar to those discussed as part of the current literature:

(1a) John sent Mother a gift.   (1b) John sent a gift to Mother.

(2a) John mailed Mother a gift.   (2b) John mailed a gift to Mother.

(3a) John found Mother a gift.   (3b) John found a gift for Mother.

(4a) John ordered Mother a gift.   (4b) John ordered a gift for Mother.

On the left, the four sentences (1a-4a) are written in a typical SVOO pattern.   The data of the language provides many examples of this input.   However, the learner must discover that these sentences can be recast with a preposition marking the indirect object.   For (1b) and (2b), the verbs 'mailed' and 'sent' take the preposition 'to.' For (3b) and (4b), the verbs 'found' and 'ordered' take 'for'.   The point of this example of dative movement is that, at some level, the rules must be organized semantically, and learners will need a certain degree of input to come to collocate 'to' with 'sent' and 'mailed,' and 'for' with 'found' and 'ordered.' Selinger (1996) makes this point by comparing dative constructions of 'give' and 'donate.'

White (1987), assuming semantic triggers must be implicit in Krashen's theory, expounds on his omission of Gleitman, Newport, and Gleitman's (1984) work about caretaker speech.   They view the here-and-now principle as unconfirmed, and conjecture the simplification of motherese may not be complex enough to effect development in child grammar.   Indeed, there is ample evidence of the controversy sur-

rounding the significance of caretaker talk in the literature. Lightfoot (1989) makes the point that immigrant children can come to acquire standard language despite the model of their parents. Schachter (1983) makes an argument against motherese in a cultural comparison of caretaker speech.

Krashen, of course, is not a UG linguist. However, his input hypothesis (Krashen, 1987) could be recast as a semantic trigger since there is a certain threshold of comprehension at which the input will initiate production. Discussion about caretaker speech may only partially reflect the development of child language since sources of input are potentially vast; motherese is only a limited part.

(d) Linguistic.

Roeper (1981) puts forth a phonological trigger. A child hearing the pronunciation of 'return' and 'rewrite' may come to differentiate the morphological function of the latter 're-,' from the former. Other linguistic considerations, such as markedness, may prove useful as explanations about mechanisms which trigger higher levels of competence as learners all progress through a similar order of development. To illustrate one interpretation of markedness, we shall consider how a child whose mother tongue is English comes to discover that subjects of verbs are systematically required.

It may be a popular assumption with regard to principle 2.4.2 (5), "Avoid Pronoun" of Chomsky's *Lectures on Government and Binding* (1993) that all children initially assume that their language, like Spanish or Italian, does not require pronouns. At the same time, inherent in the definition is that pro-drop languages do not require explicitly expressed subjects while non-pro-drop languages do not permit subject deletion. Consequently, the implication to be drawn is that pro-drop is the

unmarked, and therefore more readily accessible, parameter[4]. Let us assume, for the sake of argument, that this assumption is essentially correct. On one hand, children learning Spanish or Italian as their mother tongue assume, because of the initial pro-drop setting in UG, that they are not required to fill the subject position. On the other hand, children learning English have to somehow discover that subjects must always be provided when they arrive at the stage of standard production, prior to which utterances such as "need it?" and "got it" may have been used serviceably owing to their frequency in the input. By this time, their grammar development will have required moving the parameter from the unmarked position (+ pro-drop) to a marked one (− pro-drop). How does the child resolve this problem?

The problem is a particularly interesting one in terms of intellectual maturation since Chomsky's *Lectures* (1993) leave the reader to incorporate cognitive development. One solution is put forward by Hyams (1986). She speculates that input with empty subjects, known as pleonastics, such as "it's snowing" and "there's a toy" switch the pro-drop parameter of the child's grammar to the non-pro-drop position. The resetting of the parameter is effected through the child's realization that explicit pronouns are so prevalent that they must be present even if they refer to nothing. The result of this resetting will not only be the expected production of subjects before verbs, but also of the appearance of modal verbs, the latter being hooked into the circuit of the non-pro-drop parameter.

Hyams' explanation could be tested as a study. Consider Brown's

---

[4] This is merely one interpretation of the intent of markedness. Berwick's (1985) discussion of markedness is certainly more technical, but gives more details of the breadth of the possible intent.

(1973) mean length of utterance (MLU) as a measurement. The morphemes of an utterance were used as a measure of its length and, as an extension, a measure of children's linguistic development at various points in time using the transcriptions of their verbal production. Below a certain MLU, perhaps anglophone children exhibit a relatively consistent pro-drop use. Above another MLU, perhaps subjects occur regularly with their verbs to a degree which approximates that of adult speakers. A transcription of child-parent interaction would allow a study of the correlation between adult production and child production on this point of development at various times. Although the design of such a study appears revealing, I return to the fifth conclusion of Brown's work to discuss his caveat . He advises:

> The upshot of the several kinds of test made is that, for the 14 English grammatical morphemes, there is no evidence whatever that frequency of any sort is a significant determinant of order of acquisition (p. 409).

However, this advice is tempered in the sentence immediately following: "Still some marginal role for frequency is really guaranteed; children will not learn constructions they never hear" (p. 409). This concession in which he compares his claim to a guarantee seems to me to be an important consideration about the frequency of morphemes in the input. The frequency of the morphemes of the input could be revealing when correlated with a subject's accuracy on these points of grammar, particularly if tested at various times in the subject's language development.

Hyams' underlying framework in any case has interesting implications for developmental orders. If we envisage UG as multiple circuits with some hooked in series, and others, in parallel, then it seems

plausible that setting or resetting one parameter could simultaneously modify others. This may explain, particularly for L_2 learners, not only developmental orders, but also individual variation to some extent.

1.2.2.2 UG Framework for Explaining Transfer from L_1

One of the most useful applications for language education of the UG framework is in the area of L_1 transfer. White (1989) gathered data confirming, in this instance, the important influence of L_1 on the interlanguage development. In this study, she administered grammaticality judgment tasks to both native speakers of French and English who were learning each others' language. Consider the permutations below given the possible positions of the adverbial phrase:

(1a) I read the newspaper each morning.

(2a) Each morning, I read the newspaper.

(3a) *I read each morning the newspaper.

Standard English does not allow (3a) because of the adverbial phrase intervening between the verb and direct object. This condition is called strict adjacency.

On the other hand, the sentence could be cast in French, because of the possible positions of the adverbial phrase, in the following ways:

(1b) Je lis le journal chaque matin.

(2b) Chaque matin, je lis le journal.

(3b) Je lis chaque matin le journal.

Standard French <u>does</u> allow (3b) since various kinds of adverbials may intervene between the verb and direct object. This condition is called non-strict adjacency.

The purpose of White's study was to see whether francophones

would reject adverbial phrases intervening between the verb and direct object in English like example (3a). Similarly, White wished to determine the extent to which anglophones learning French would accept constructions such as (3b). She found that, for the most part, the L_2 learners stuck with the parametric setting of their mother tongues. Anglophones, for the most part, rejected constructions like (3b). The few who did not, however, seemed to be aware of what was appropriate as an intervening adverbial. At the same time, francophones learning English judged constructions like (3a) acceptable.

From her observations, White concluded that: (1) transfer can be attributed to the L_1 parameter which the learner does not easily abandon; (2) fossilization can be explained as the result of not resetting a parameter to an L_2 position when the new position allows only an L_1 sub-set. (In this study, English permits a subset of those positions permitted in French for adverbials); (3) Subtle UG influences permit those English-speaking learners who have reset the adjacency parameter to access other rules about what adverbials are permissible in the intervention between French verbs and their objects.

A logical extension of her third conclusion might be that UG consists of "circuits" joining the parameters. Consequently, switching one parameter could simultaneously activate access to other rules, or even the resetting of other parameters. In this way, as White (1985b) claims, we can provide a framework which reconciles such dichotomies as contrastive analysis versus creative construction (Dulay, Burt, & Krashen, 1982).

2. THE EXPLICIT/IMPLICIT OPTION

The second issue which Stern raises is the explicit/implicit option. This refers to the controversy surrounding the learning of the target

language "rules." On the contrary, perhaps conditions which favour picking up the language to be acquired can be recreated in the classroom.

2.1 Analytical or Intuitive Learning

Through the explicit option, language can be studied analytically through descriptions of structure. Through the implicit option, language can be acquired intuitively, sometimes naturalistically. The most typical application of the first approach is formally studying grammar, vocabulary, and phonetics. The justification for this kind of study is made on cognitive grounds. Adherents of explicit methods, for the most part, see language learning as an intellectual process. Stern names Carton (1971) as a proponent from this school of thought.

As for intuitive approaches, there seem to be two interpretations as to what constitutes learning intuitively. Rivers (1964), one of the major adherents of audiolingualism, claimed that teaching through analogy was preferable to analysis. Her claim could have been endorsed as well by audiovisual and certain direct method proponents. Showing the learners what possibilities were appropriate in various paradigms and having them develop good habits through drill reflected an approach to learning consistent with behaviourism.

More recently, communicative language teaching (CLT) is generally the expected classroom standard. Its approach to learning would also be considered one that appeals to intuitive or natural language processes. However, the method realized by this approach is considerably different from its audiolingual precursor. Most CLT advocates would stress the importance of "naturalistic" input, which they liken to motherese. In this way, they presume to effect competence through an acquisition similar to that of the first language.

Proponents of various methods have typically treated the explicit/ implicit option dichotomously. In fact, bipolarization is probably too simplistic. Many other factors need be considered in addition to the underlying approach to learning. Cognition, age and learning styles are such variables. We cannot expect a highly analytical well-educated adult to learn in the same way as a pre-literate child. Hence, we expect very different classroom procedures even if the approaches to learning and to language are the same. Whereas the adult in this hypothetical situation will probably benefit from explicit presentations, the child will probably not.

2.2 UG and Models of Language Learning

Where a UG framework can benefit the second language educator is by making our models of learning and language more complete. Although Krashen's distinction between learning and acquisition has recently gained the greatest prominence, similar distinctions have long been made. Other recent second language researchers who have posited models on an implicit/explicit distinction include Macnamara (1973), McLaughlin (1978), and Bialystok (1978).

The framework of these descriptions are "black boxes." In the case of Krashen and Bialystok, there are schematic representations with relationships implied between the constituent parts. However, how can the contents be described? How do the parts work together? The whole represented by the model will no doubt be greater than the sum of theories, both language and learning, which it represents.

Take the well-known Bialystok model for instance. Central to the model are three boxes representing kinds of knowledge: implicit linguistic knowledge, explicit linguistic knowledge, and other knowledge. The two former require some theory of language. In such cases, UG

can give us some idea about what the content of the black box is.

Elaborating models of language learning through a more exact description of UG is a highly desirable goal in order to comprehend the interaction between the constituent parts of the diagram.   At the same time, succinct applications of language theory will elucidate points of comparison for teachers and their students.

2.3 Error Correction and the Subset Principle

In the meantime, we can begin to glean certain considerations from the UG framework to guide us in our teaching procedures.   With regard to the controversy about what should be taught explicitly, we can be guided by considering parametric variations.   In particular, we can make a reasonable explanation about the nature of certain errors in the interlanguage.   At the same time, through parametric considerations, we can speculate about whether input (implicit learning) or instruction (explicit learning) would be more likely to eradicate the errors.   Of course, the linguistic "prescription" may be over ridden by various factors such as level of cognition, learning styles, or lack of metalinguistic awareness.

As an example of the extent to which we may find parametric settings useful in determining how to correct, let us consider how the pro-drop parameter can be reset for L_2 learners of English and Spanish. In this example, English, in comparison to Spanish, allows a subset of subject phrases.   Spanish, allowing certain implicit (i. e., unstated) subject phrases which English prohibits, represents the superset.

White (1990) has hypothesized that when parameters are related as superset and subsets:

1. Positive evidence (i. e. implicit learning) alone may be sufficient for switching the parameter from the subset to the superset;

2. Negative evidence (i. e. explicit learning) will be required to switch the parameter from the superset to the subset.

Consequently, anglophones learning Spanish will discover that the subject does not need to be expressed explicitly through input with implicit subjects. On the other hand, Spanish-speakers, believing English is like their mother tongue, will not discover from mere input by itself that their production is ungrammatical. The data will not seem dissonant with the parametric setting of the interlanguage. To expunge utterances like "*Speak Spanish very well" (when the subject I is intended rather than a command), Hispanic learners need to have their attention drawn metalinguistically to overcome the pattern attributable to their mother tongue.

2.4 Quantity of Positive Evidence

Although I believe White's contention to be essentially correct, perhaps it need be nuanced to some extent. Even assuming that positive evidence can effect change in the learner's grammar, White's rule-addition hypothesis does not imply a "free ride" for second language learners or teachers. Let us return to White's grammaticality judgment study of the adjacency principle (1989). In this study, anglophones held tenaciously to the L_1 adjacency parameter even though positive evidence could have theoretically triggered access to a "looser" grammar. After the fact, it is difficult to determine the amount of positive evidence that these learners received, but my intuition suggests that they likely received very little. As in this sentence, the intervening adverbial s'emploie surtout pour étonner, avec une certaine intensité, l'interlocuteur. The surprise effect is achieved by keeping the interlocutor in suspense for the object. There are, of course, other

reasons for intervening adverbials in French. However, I suspect most often the sentence is cast in this manner for stylistic effect. Furthermore, it is a style more appropriate for written than spoken language. Considering the special status of intervening adverbials, we expect that FSL teacher talk or modified-French-for-anglo-Montrealers is an impoverished source of this kind of evidence.

I believe White's positive evidence speculation needs to be complemented with some mechanism to account for the frequency of the data to the teacher. Positive evidence by itself may effect restructuring, but below a minimal threshold level, positive evidence is unlikely to effect change in the interlanguage. Whereas anglophones learning Spanish are likely to begin deleting pronouns relatively early, the same learners may take considerably longer to abandon strict adjacency in French. On one hand, Spanish data will provide abundant evidence of pro-drop; on the other hand, French data will provide relatively scanty evidence of non-strict adjacency. Although both changes are predicted by White's speculation about supersets, the success rate on acquiring the superset may be attributed to the frequency of the positive evidence.

2.5 Pedagogical Application

The subset hypothesis coupled with what could be termed the "frequency factor" could guide the teacher in determining what is taught explicitly or implicitly.

First, the resetting of a parameter from a superset to a subset will probably require error correction, at a minimum. Perhaps the learner will eventually come to infer the rule through systematic correction. This teaching strategy will more readily apply to intuitive learning styles or to learners who lack sufficient cognizance or metalinguistic awareness. On the other hand, we should not deprive literate schooled

adults of an analytical presentation. We would expect their cognitive development to be sufficiently advanced for them to benefit from an explicit metalinguistic explanation.

As for the resetting of a rule-addition parameter, this adjustment will also require some consideration of the "frequency factor." If the language is inherently full of positive evidence, the learning should proceed implicitly with input alone motivating the resetting of the parameter. I have speculated that anglophones acquiring Spanish pro-drop is an example of this phenomenon.

In contrast, if the target language offers infrequent positive evidence, the input will require some treatment to make the feature more salient. To this end, perhaps the implicit teaching technique of the audiolingual method would prove useful. Anglophones giving up the strict adjacency principle would be possibly more inclined if presented with paradigms which make salient the various positions of French adverbials. This would require some manipulation of the input in order to make the positive evidence more apparent.

Even though the presentation of paradigms may be, to some extent, inconsistent with "naturalistic" input hypotheses, such a technique would be consistent with Rutherford and Sharwood Smith's (1985) contention that conscious raising "can have degrees (their emphasis) of explicitness" (p. 275). This is an axiom to Sharwood Smith's (1980) pedagogical grammar hypothesis which predicts that, given certain circumstances, explicit learning focused on "structural regularities" can effect a greater "rate of acquisition" than can naturalistic learning[5].

---

[5] Readers looking for the "pedagogical grammar hypothesis" will find it in Rutherford and Sharwood Smith's (1985) article, more readily accessible than Sharwood Smith's (1980) conference presentation.

Through this discussion of error correction, I have dealt critically with the subset principle. It is, nonetheless, an essential index for foreign and second language practitioners. It predicts fossilization and gives some indication about what should be taught implicitly and explicitly. In addition, it provides specific targets for error correction, which is generally not set out descriptively in an incremental fashion. We should receive this framework for error correction enthusiastically. Because of it, we can avoid the frustration of trying to correct everything sporadically, or in the absence of a compass for direction, relying heavily on input alone.

3. THE CODE-COMMUNICATION DILEMMA

The third issue identified by Stern, the code-communication dilemma, reflects the controversy related to whether a second language is best taught in a situation which requires learners to focus on how the language works, or instead, in one which allows them to learn through the experience of tasks requiring communication in the target language. Classroom teaching has generally proceeded in the first way. Either through analysis or analogy, learners are to come to a better understanding about how the language works as a system.

The "communication" alternative is exemplified by the French immersion programs in Canada. In such programs, the target language is used as the means of instruction, and the children focus their attention on the content and information, rather than on how the language works. In this way, immersion programs are intended to exploit the early-life interaction of children who are acquiring the first language. The underlying assumption is that these same dynamics can effect acquisition in the classroom for the immersion pupil without hampering L_1 literacy skills (Genesee, 1987).

Whether teaching communication or teaching code is more desirable cannot be determined through a theory of language by itself. What is most appropriate will again depend on a number of variables. Richards and Rodgers (1986) have elaborated a trichotomy which describes the role of learning and linguistics in teaching methods. At the top of the hierarchy is the approach, which is a consideration of theories of language and learning. At times, a theory of language drives the approach. In such methods, theories of learning are subordinated to linguistic considerations. Other times, a theory of learning predominates. In these methods, greater consideration is accorded to such factors as group dynamics, learning styles, learner needs, and motivation than to the theory of language. An approach slanted towards learning which is realized through activities focusing on communication seems logically more compatible for students with lower levels of cognitive development and intuitive learning styles. And although communication will necessarily be the terminal objective for highly schooled analytical learners, perhaps they will reach this end more efficiently in the classroom through an analytical linguistic approach.

3.1 UG and a Specific Linguistic Approach

3.1.1 Pedagogical Grammars

While UG cannot help us to determine whether a given student will learn more through the code or through communication, UG will help the teacher determine the path to follow. In the first place, let us consider a linguistic approach. The first need will be pedagogical grammars which have sorted the structures of the target language according to the principle of simple to complex. The grammars should ideally take into consideration what orders of difficulty are inherent in

the language, and what further complication can be expected from mother tongue interference.

Sharwood Smith (1981) has expressed a useful distinction for pedagogical grammars. He differentiates concentrated grammars from extended grammars. Whereas the former are intended as general references for the teacher and advanced learner, the latter are aimed at specific linguistic backgrounds, and are therefore closer to these learners' most immediate needs.

The content of both kinds of grammars must be intimately linked to a theory of language. In the case of a concentrated grammar, consideration need be given to such things as language universals and common developmental orders. In the case of extended grammars, to L_1 transfer. UG can therefore prove useful. In the consideration of language universals and common developmental orders, grammarians can turn to such principles as markedness. As for L_1 transfer, they can be guided by such indices as the subset principle.

3.1.2 Linguistic Syllabi

Once the pedagogical grammars are in place, we need to select those aspects of the target language which we can reasonably expect to teach during the allotted period of instruction. Pienemann (1985) proposes organizing syllabi according to the "learnability/teachability" order. He posits that there is an inherent progression in language development determined by the learner's internal syllabus. A learner may progress from stage X to X+1, but will be unable to control the structures of X+2 without first mastering those of X+1.

Pienemann has begun to describe the structures of five learnability stages for German based on data gathered from both tutored and untutored learners. (See, for example, Pienemann, 1987). Although he

is able to describe the observable stages for Romance learners of German with modest detail, his learnability hypothesis would have greater generalizability to other L_1's and L_2's if we were able to describe the underlying mechanism at each stage. To this end, a more sophisticated model of UG could serve our needs. What principles can we use to explain the inherent simplicity for stage X over X+1? subsets versus supersets? degrees of markedness? the emergence of certain constraints?

If we can construct a universal learnability hierarchy based on universal principles of UG then we should be able to predict the expected learnability orders for specific L_1 and L_2 combinations. This learnability hierarchy would be a useful reference for determining the linguistic content of any second language course.

3.2 UG and Learning-Based Approaches

The benefit of a theory like UG seems undisputable to practitioners when the approach subordinates a theory of learning to a theory of language. In this case, a highly elaborated UG theory will pave the way for the syllabus designer.

On the other hand, when the theory of learning predominates, a theory such as UG could occupy a less apparent role, but an important one, nonetheless. In such cases, the theory of language will provide sign posts for the syllabus designer. Assume, for instance, that the syllabus is organized around a certain number of functions. One objective is expressing speculation. A prerequisite skill would be some minimal level of control of modal verbs. In order to determine an opportune time to slot "expressing speculation" in the syllabus, we would ideally need to know the amount of naturalistic classroom input required to reset the parameters which permit adequate control of the

modal verbs. Although such an elaborate UG model would require time to develop, we could, for the time being, grade syllabus content more confidently by considering both developmental and learnability orders described in current literature.

To conclude the discussion of how a theory of language can serve CLT, let us consider communicative competence. This is presumably the competence educators wish to effect through the communicative approach. Canale and Swain (1980) originally named three components to make up communicative competence: (1) grammatical competence, (2) sociolinguistic competence, and (3) strategic competence. Under the heading of "grammatical competence," Canale and Swain express reluctance to adhere to any linguistic theory although they acknowledge the significance of grammatical accuracy for expressing and interpreting "literal meaning" (p. 30).

On one hand, communicative competence assumes some logical linguistic rationale. Yet, on the other hand, the original proponents of communicative competence are reluctant to endorse a current linguistic theory. This is an unfortunate gap, and requires a theory like UG to turn a hollow shell into a substantive framework. This gap is also a loophole for not explaining well enough the regularities which can be presented to learners about grammar, vocabulary, and even phonology.

3.3　A Case for Testing

A familiar adage confirms that the proof is in the eating. I will conclude my paper by reporting briefly on some testing I have conducted using a UG instrument (Kirkwold, 2004) as a measure of proficiency. I used the test described by White (1985a) and administered it to 14 classes in three universities, including our own, in the Sapporo region when I had begun to teach part-time at Otaru University of Commerce

(OUC). My ranking resulted in an order which could be confirmed by similar measures of proficiency.

When I administered this testing in 2001, I did not ask OUC students for their majors. However, in subsequent years, their majors have always included law, economics, and commerce. I suspect these majors would also have been indicated in 2001. I was satisfied to confirm through this test that the group averages for the two classes (one a first-year day class, the other a second-year night class) I was teaching at OUC were not as high as for first-year classes at Hokkaido University (HU), with only one exception to this trend. First-year students at HU whose majors were agriculture/veterinary, medical/dental, and engineering had higher group averages than did the first-year day class at OUC. HU education/law students had a higher group average than did the second-year night class at OUC, but it was lower than the OUC first-year students. It is the comparison of the OUC day students with the HU education/law students that makes the one exception. I used these results to reset the focus in the courses I was teaching at OUC. Whereas the textbook I had chosen for the OUC students was intended for the revision and practicing of problematic sounds of spoken English, I chose other materials to use in addition to make these courses more similar communicatively by way of approach to the instruction I provided during class time in Sapporo.

Qualitatively, the ranking I have just described resonates on the whole with my experience teaching these students over the years. Descriptively, the ranking appeals to what could be expected for students in their majors and institutions. The order also suggests more specific questions as to differences between the same majors at various universities. For example, given any two classes of law majors at Japanese universities, would significant differences be revealed when

this instrument is used as a test of proficiency?

From my own experience, I would endorse this UG instrument as a measure of proficiency. I have used the test at the beginning of subsequent courses to compare group scores. The testing done at an early date is an indicator to me of how much the students taking a given course have changed since the previous term. A class average may show variability from year to year and the extent to which the rank order of the groups has achieved relative stability over the years could be documented as confirmation of a trend.

As for the selection of a UG instrument, a classroom teacher may find the related literature useful in understanding problematic points of grammar. As a teacher of writing, I am left to consider the relevance of the pro-drop parameter as it is applicable to Japanese learners of English. I am curious to learn if there are properties of this parameter associated with persistent mistakes in student writing. Given that the pronoun is not necessarily expressed in Japanese, this potentially is a source of negative transfer. I recall teaching ESL in Canada that Hispanic writers of English would occasionally omit the subject pronoun, which seems consistent with the pro-drop designation for Spanish. In eleven years of teaching in Japan, however, the number of times I have encountered this omission has not made an important impression on me. Reflecting on the suffixation of the Japanese verb system, I have determined, I believe, why omitted pronouns are unlikely candidates for negative transfer. Consider the present tense in English. It is only the third person singular which requires the-s morpheme. Native speakers of English do not therefore train themselves to infer the subject from the absence of suffixation for all but the third person singular. They must listen instead for a subject or pronoun in its place. Speakers of Spanish and Italian, however, may infer subjects from the

suffixation of their verb systems since inflectional agreement for verb endings is much more complex than the English one. Japanese is more like English since the morphology is such that the suffix indicates tense (either present or past) but not person. The same suffix is applied irrespective of subject (and therefore grammatical person). In the usual discourse of their language, Japanese speakers would consequently not expect to infer subjects from verb suffixation. Although the subject may be omitted, it would necessarily be expressed when the intent cannot be inferred. In spoken Japanese, there would be confirmation or clarification with the interlocutor. In written Japanese, the occurrence of the pronoun or subject would therefore be present frequently to avoid ambiguity. This, I surmise, is the explanation for pronouns being rarely omitted by Japanese learners writing English.

That Japanese, Spanish, and Italian are all pro-drop languages does not make them similar with regard to what is inferred from or implied by the unexpressed subject. The suffixation of the verb reflects the grammatical person of the subject for the two Romance languages. That an interlocutor can imply this much in the discourse of these languages probably results in the initial transfer of omitted subjects when native speakers of Spanish and Italian write in English.

As for Japanese, who can imply tense through verb suffixation, they must confirm with their interlocutors in spoken discourse when the implied subject is misunderstood. From their experience in conversation, they must subsequently realize that writing their own language will require subjects be included deliberately to avoid ambiguity. When they learn to write English, this same expectation will transfer, and they will not make the pro-drop deletions attributable to Spanish and Italian learners.

I have explained why pro-drop does not create the seemingly

obvious problem associated with the name; however, there may be other problems for EFL teachers in Japan which parameters may explain. The frequency of 'because' clauses left unsubordinated suggests to me that further investigation of parameters may be a possible avenue to understand the occurrence of this common mistake by Japanese students in writing courses. By going beyond the surface structure, perhaps teachers could discover the appeal of unsubordinated 'because' clauses through arrangements of subsets. At the same time, the frequency that could not be justified stylistically of 'and,' 'but,' and 'so' occurring at the beginning of sentences is a second problematic point which is worthy of similar investigation.

REFERENCES

Andersen, R. (1977). The impoverished state of cross-sectional morpheme acquisition/accuracy methodology. *Working Papers on Bilingualism, 14,* 47-82.

Andersen, R. (1979). The relationship between first language transfer and second language overgeneralization: Data from the English of Spanish speakers. In R. W. Andersen (Ed.), *The acquisition and use of Spanish and English as first and second languages* (pp. 43-58). Washington, DC: Teachers of English to Speakers of Other Languages (TESOL).

Archibald, J. & Libben, G. (1995). *Research perspectives on second language acquisition.* Toronto: Copp Clark.

Asher, J. (2000). *Learning another language through actions* (6th ed.). Los Gatos, CA: Sky Oaks Productions, Inc.

Bailey, N., Madden, C. & Krashen, S. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning, 24,* 235-243.

Berwick, R. (1985). *The acquisition of syntactic knowledge.* Cambridge, MA: The MIT Press.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28,* 69-83.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Carton, A. (1971). Inferencing: A process in using and learning language. In P. Pimsleur & T. Quinn (Eds.), *The psychology of second language learning* (pp. 45-58). Cambridge University Press.

Chomsky, N. (1993). *Lectures on government and binding: The Pisa lectures*. New York: Mouton de Gruyter.

Chomsky, N. (1998). *On language*. New York: The New Press.

de Villiers, J. & de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*, 267-278.

Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*, 37-53.

Dulay, H., Burt, M. & Krashen, S. (1982). *Language Two*. New York: Oxford University Press.

Felix, S. (1984). Maturational aspects of universal grammar. In A. Davies, C. Criper, & A. Howatt (Eds.), *Interlanguage* (pp. 133-161). Edinburgh University Press.

Genesee, F. (1987). *Learning through two languages*. Boston, MA: Heinle & Heinle.

Gleitman, L., Newport, E. & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11*, 43-79.

Greenbaum, S. & Quirk, R. (1990). *A student's grammar of the English language*. Essex: Longman Group.

Hakuta, K. (1976). A case study of a Japanese child learning English as a second language. *Language Learning, 26*, 321-351.

Huebner, T. (1979). Order-of-acquisition vs. dynamic paradigm: A comparison of method in interlanguage research. *TESOL Quarterly, 13*, 21-28.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Tokyo: D. Reidel Publishing Company.

Kirkwold, L. (2004). Issues in course development and testing: An Eigo

course evaluation at OUC. *Language Studies, 12*, 15-21. Otaru: Otaru University of Commerce.

Krashen, S. (1987). *Principles and practice in second language acquisition.* New York: Prentice-Hall International.

Krashen, S. & Terrell, T. (1998). *The natural approach: Language acquisition in the classroom.* Hertfordshire, UK: Prentice Hall Europe.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: The University of Michigan Press.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409-419.

Larsen-Freeman, D. & Long, M. (1991). *An introduction to second language acquisition reserach.* Essex, UK: Longman Group UK Limited.

Lightfoot, D. (1989). The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences, 12*, 321-334.

Macnamara, J. (1973). The cognitive strategies of language learning. In J. Oller, Jr. & J. Richards (Eds.), *Focus on the learner: Pragmatic perspectives for the language teacher* (pp. 57-65). Rowley, MA: Newbury House.

McLaughlin, B. (1978). The monitor model: Some methodological considerations. *Language Learning, 28*, 309-332.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 23-75). Clevedon, UK: Multilingual Matters Ltd.

Pienemann, M. (1987). Determining the influence of instruction on L_2 speech processing. *Australian Review of Applied Linguistics, 10*, 83-113.

Pinker, S. (1995). *The language instinct.* New York: HarperPerennial.

Richards, J. & Rodgers, T. (1986). *Approaches and methods in language teaching: A description and analysis.* Cambridge University Press.

Rivers, W. (1964). *The psychologist and the foreign-language teacher.* University of Chicago Press.

Roeper, T. (1981). On the deductive model and the acquisition of productive morphology. In C. Baker & J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 129-150). Cambridge, MA: The MIT Press.

Rutherford, W. & Sharwood Smith, M. (1985). Consciousness-raising and universal grammar, *Applied Linguistics, 6*, 274-82.

Schachter, J. (1983). Nutritional needs of language learners. In M. Clarke & J. Handscombe (Eds.), *On TESOL '82: Pacific perspectives on language learning and teaching* (pp. 175-189). Washington, DC: Teachers of English to Speakers of Other Languages.

Seliger, H. (1996). Primary language attrition in the context of bilingualism. In W. Ritchie & T. K. Bhatia (Eds.). *Handbook of second language acquisition* (pp. 605-626). Tokyo: Academic Press.

Sharwood Smith, M. (1980, December). The competence-performance distinction in the theory of second language and the pedagogical grammar hypothesis. Paper presented at the Contrastive Linguistics Conference, Boszkowo.

Sharwood Smith, M. (1981). Notions and functions in a contrastive pedagogical grammar. In A. James & P. Westney (Eds.), *New linguistic impulses in foreign language teaching* (pp. 39-53). Tübingen, Germany: Gunter Narr Verlag.

Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford University Press.

White, L. (1985a). The "pro-drop" parameter in adult second language acquisition. *Language Learning, 35*, 47-62.

White, L. (1985b). Universal grammar as a source of explanation in second language acquisition. In B. Wheatley, A. Hastings, F. Eckman, L. Bell, G. Krukar, & R. Rutkowski (Eds.), *Current approaches to second language acquisition: Proceedings of the 1984 University of Wisconsin-Milwaukee Linguistics Symposium* (pp. 43-68). Bloomington, IN: Indiana University Linguistics Club.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics, 8*, 95-110.

White, L. (1989). The adjacency condition on case assignment: Do L_2 learners observe the subset principle? In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 134-158). Cambridge University Press.

White, L. (1990). Implications of learnability theories for second language learning and teaching. In M. A. K. Halliday, J. Gibbons, & H. Nicholas

(Series Eds.), *Learning, keeping and using language: Volume 1* (pp. 271-286). Amsterdam: John Benjamins Publishing Company.