OCCASIONAL PAPER #12
1985

SYNTACTICIZATION OF TOPIC
IN JAPANESE AND MANDARIN STUDENTS'
ENGLISH: A TEST OF RUTHERFORD'S MODEL

Patricia Ann Duff

DEPARTMENT OF ENGLISH *AS* A SECOND **LANGUAGE**

UNIVERSITY OF HAWAII AT **MANOA**

O C C A S I O N A L   P A P E R   S E R I E S

In recent years, a number of graduate students in the Department of English as a Second Language have selected the thesis **option** as part of their Master of Arts degree **program.** Their research has covered a wide range of areas in second language learning and teaching. Many of these studies have attracted interest from others in the **field,** and in order to make these theses more widely available, selected titles are now published in the Occasional Paper Series. This series, a supplement to the departmental publication Working **Papers,** may also include reports of research by members of the ESL faculty. Publication of the Occasional Paper **Series** is underwritten by a **grant** from the Ruth Crymes Scholarship Fund. A list of available titles and prices may be obtained from the department and is also included in each issue of Working Papers.

The reports published in the Occasional Paper Series have the status of **"progress** reports", and may be published elsewhere in revised form.

Occasional **Paper** #12 is an **MA** thesis by Patricia Ann Duff. Her thesis **committee** members were Craig Chaudron (chair), Jean **Gibson,** and Richard Schmidt. This work should be cited as follows:

DUFF, Patricia Ann. 1985. Syntacticization of Topic in Japanese and Mandarin **Students'** English: A Test of **Rutherford's** Model. Occasional **Paper** #12, Honolulu: Department of English as a Second Language, **University** of Hawaii at **Manoa.**

ABSTRACT

Rutherford (1983) drafted a two-part model to account for the syntacticization of Topic in the English of Japanese and Mandarin learners. For Japanese, he charted the acquisition of English existential constructions with there out of earlier topicalized locative expressions. He characterized Mandarin learners[1] development in terms of the evolution of Subjects from earlier existentials and Topic-Comment constructions.

Implicit in Rutherford[1]s model are assumptions concerning (1) the role of transfer in second language acquisition; (2) typological distinctions between English, Japanese, and Mandarin based on the roles of Topic and Subject; and (3) the naturalness of the developmental shift from Topic Prominence to Subject Prominence. This theoretical background was reviewed to provide rationale for Rutherford's claims and motivation for the hypotheses tested in the present study.

Written compositions of 105 Japanese (J) and 105 Mandarin (M) learners[I] whose proficiency ranged from TOEFL 450-599, were examined. Analysis of variance was used to determine the effect of the independent variables of first language and proficiency on the dependent, variable of syntacticization of Topic.

In general[r] Rutherford's model was not supported by statistically significant results[r] although the data revealed trends in the predicted direction for most measures. The study did, however[r] provide statistical support for differences between language groups in the production of passives $(J > M)$, Subject-verb agreement $(J > M)$, PRO-drop $(M > J)$, and serial verbs $(M > J)$; proficiency did not have a significant main effect in all of these cases though.

The results could be explained[I] in part, by typological differences between Japanese and Mandarin. However, it was noted that the research questions might be addressed more satisfactorily by conducting further studies with learners at lower levels of proficiency[I] and by examining oral production data, in addition to written data.

iii

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

. . . Recently when I was talking to a friend, I attempted to say in English that a large proportion of the population of Hawaii is of Asian extraction. My first utterance was "Hawaii . . .," because of transfer from my first language, Japanese. That is, in Japanese, a topicalized noun is likely to be placed at the beginning of a sentence, like

hawaii wa ajia-jin ga ooi
Hawaii TOP Asians SUBJ many
'Regarding Hawaii, Asian people are many'=
There are many Asians in Hawaii

This topic-initialization tendency is transferred into my English interlanguage.
Instantly, I recognized that I could not go on to constitute a grammatical sentence, because English does not have a specific syntactic/lexical device (e.g., postpositional particle) that marks a sentential topic, and neither does it allow a topical noun phrase without case. I thus was obliged to restate the proposition, changing "Hawaii* into the locative "In Hawaii, . . ." and I finished by saying, ". . . there are many Asians." I did not think of marking "Hawaii* for a locative case initially, because in the corresponding -Japanese sentence "Hawaii" is not marked for case. In a word, when I started verbalizing, I did not have a clear idea what sort of sentence I was going to utter and just followed a typical Japanese strategy of case marking and topic initialization (Sasaki 1985:8). [note 1]

1.1 Overview

When certain patterns of linguistic behavior occur sufficiently frequently in second language acquisition data, researchers attempt to determine their possible sources with-some explanatory power: that is, whether they are (1) strictly developmental (in a non-biological sense), in that they are shared by learners from a variety of

language backgrounds and are second language (L2)-specific (and are thus a result of the language structure in the input); (2) whether they are a consequence of linguistic structures and parameters in the first language (L1), which are transferred to the L2, and are thus L1-specific; (3) whether they are a product of universal grammatical principles of natural languages, which are attested in developmental studies of both L1 and L2 acquisition (SLA), diachronic language change, pidginization and creolization; or (4) whether they are a combination or interaction of the above and other factors, such as psycholinguistic processing constraints and markedness theory.

A question that has long been asked by and of teachers and researchers alike is related to the second point above, namely, How do learners' L1s affect their second language acquisition (SLA)? In recent years, a small subset of applied linguists who have again been grappling with this topic have come to recognize language typological features as a potential source of transfer, especially at the discourse-syntactic and pragmatic levels (e.g., Huebner 1979, 1982, 1983, 1985, Gass and Selinker 1983, Givon 1979, 1984, Rutherford 1983, Eckman et al. 1904, Hilles 1985, White 1985). As Rutherford (1984b:146) notes,

> The point of departure for typologically oriented L2 transfer research is of course the fundamental assumption that the different formal shapes of IL syntax will be traceable in part, and in varying degrees, to the intersection of L1 and L2 typologies.

Researchers observe that aspects of global typological organization in the L1 can be unwittingly carried over into SLA, partly due to the compatibility of pragmatic strategies with other constraining factors in language development, such as those listed above. Therefore, in addressing the age-old research question concerning transfer, it is now apparent that researchers should also examine the degree to which transfer of a given typological parameter is natural, unmarked, productive, and consistent with other principles of language change. Secondly, in conducting research of this type, they explore the direct and indirect linguistic reflexes of the transfer effect, in terms of L2 forms which are adopted to fulfill the discourse-syntactic functions assumed by other structures in the L1.

It is thus claimed that L1 discourse-level syntax, as determined by language typology, in particular, constrains

SLA in possibly a more far-reaching and persistent manner than surface syntax alone does. Since each typological parameter is usually associated with a clustering of structural and functional properties, it is potentially a more powerful source of transfer, and thus also provides a more powerful explanation for a number of transfer effects heretofore considered to be unrelated.

For example, transfer of pragmatic strategies based on Topic Prominence in the L1 can be reflected in the IL production of Topic-Comment constructions, often without Subjects or verb agreement. Thus ESL learners fall back on strategies developed in the L1 to introduce given (old, presupposed) referents and to distinguish them from assertions and new referents* even though the L2 has different grammatical devices to convey the same pragmatic information. Hence, learners whose L1 discourse-level syntax has grammaticized Topics (e.g., Japanese) tend to use marked Topic-Comment constructions (TCs) in ESL to make Topics salient in ways that are possible in the L1. Or in languages whose word order signals pragmatic information (e.g., Mandarin, Hmong), such that preverbal information automatically has definite reference, an IL paradigm of ESL article usage lacking definite articles for Subjects is evidence of another functional transfer effect, in this case of pragmatic word order.

It is understandable that all L2 learners might at some point employ topicalization or article deletion as a pragmatic means of simplifying syntax and facilitating communication. However, the consistency and persistence of the TC usage in the ESL of native speakers of languages which are Topic Prominent is said to be characteristic of those learners' production, compared with the production of native speakers of more Subject Prominent languages.

In the ESL of Japanese (J) and Mandarin (M) students, for example, several studies report a tendency for M speakers to use an abundance of Topic-Comment (TC) constructions which, at the earliest stages, do not reflect or match the basic Subject-Predicate structure of English sentences; this interlanguage (IL) feature is also claimed for J learners (Schachter and Rutherford 1979, Rutherford 1983), but to a lesser extent. The same effect of Topic Prominence is not seen in the ILs of speakers from other language groups, though, as determined by comparisons with Arabic and Spanish data. It is argued that when Oriental learners of English (e.g., Hmong, Korean, J, M) produce IL sentences of a primarily TC nature, this can be attributed to the fact that the L1s are, generally

speaking, more 'Topic Prominent[n] than 'Subject Prominent,' according to Li and Thompson's typological classification (1976).

Not only is there an apparent effect for Ll type, there appears also to be an effect for L2 proficiency on the degree to which an IL will be moulded by the Ll-based strategies. Thus, in research of this nature, it is useful to consider the diachronic syntacticization process of discourse-level syntactic features, as in longitudinal case studies (e.g., Euebner 1979) or, alternately" to look at cross-sectional data of learners at various levels of L2 proficiency. The study which this thesis principally pursues, and which will be presented below, is a cross-sectional analysis of the development toward Subject Prominence by J and M learners of ESL.

1.2 Rutherford's Model

Rutherford (1983) examined written data of ESL learners in an American university context, and was able to derive more specific conclusions about the J and M transfer effect. He drafted a tentative but testable model to conceptualize the syntacticization processes which are representative of J and M learners! but also,. which reflect certain differences between the two groups as-well.

To summarize his findings, at the lowest proficiency levels represented in his data, M learners (whose Ll is of the attested 'Topic Prominent[n] type) produce many constructions that are fundamentally TCs, of the type typically found in Mandarin Chinese (Li and Thompson 1976, 1981, Barry 1975, Xu and Langendoen 1985); these TCs often have unexpressed Subjects and little verbal morphology. Later, Subject is produced in surface structure, following the Topic, as in left-dislocations and 'double Subject[n] constructions. Next, existential constructions with there evolve and these are used in various ways (apparently reflecting different proficiency levels), to introduce new Topics. The existentials develop progressively, along with infinitival complements and relative clauses. In the last stages, the function of there as a syntacticized dummy Subject with a grammatical predicate is learned, and finally, the Topic and Subject functions converge.

J learners, however, (whose Ll is of the dual 'Topic Prominent and Subject Prominent[n] type) also topicalize elements quite freely at the lower levels, but, Rutherford claims, there are both qualitative and quantitative

4

differences in the way J and M learners present topical material. The quantitative difference is in terms of the sheer number of total Topic-Comment constructions used by the two groups; M learners produce more TCs than J. Qualitatively, however, Rutherford notes that J tend characteristically to introduce Topics with locative expressions, such as in sentence-initial prepositional phrases with the preposition in. There is, at the same time, some overlap in the M and J IL usage in the sense that J also use existentials in a way that reflects their L2 development, but most typically they occur together with a locative expression.

Another reported difference is that J produce more extraposition with it than M, and the construction is said to introduce future Topics. This tendency for J to "overproduce" (or "overgeneralize") dummy Subjects in ESL, both it and there pronouns, has been attributed to the notion that J language structure has word order with both grammatical and pragmatic functions, whereas Chinese reserves word order for pragmatic purposes (Thompson 1978, Rutherford 1983). Rutherford thus posits differences in L1 typological parameters, such as Topic Prominence and Grammatical/Pragmatic word Order, as a way of explaining IL differences between the two groups.

1.3 The Task of Syntacticizing Topics in SLA

The task faced by these ESL learners whose L1s are not Subject Prominent is to gradually acquire the English Subject-Predicate configuration by "syntacticizing" the L1 base structure (Givon 1979); this is achieved when the behavioral properties (e.g., deletion, movement, control of coreference) and coding properties (e.g., position, case-marking, agreement) of English Subjects (cf. Cole et al. 1980, Keenan 1976) are mastered. One of the obvious properties of English Subjects is that, unlike Subjects in PRO-drop languages (e.g., J, M, Spanish), they must be overtly expressed in surface syntax (see White 1984, 1985, Hilles 1985, Zobl 1985b, Duff 1985, concerning the PD parameter in second language acquisition). The most obvious coding property of English Subjects is the Subject-verb agreement concerning which Givon (1979:209-210) writes:

> one of the most acclaimed properties of subjects, that of, arammatical aareement on the verb, is fundamentally a topic property, and . . . it arises diachronically via the reanalysis of topic

into subject and --simultaneously-- of an anaphoric pronoun into a (normally verb-bound) agreement morpheme . . . One must remember, however, that English has both SUBJECT and TOPIC constructions, and that they serve normally different discourse functions. Thus the grammaticalization of topics into subjects does not mean that the language has lost the topic construction, but only that it has gained grammatical agreement as an added morphological coding property for its grammatical subject.

Givon notes that in this process of syntacticization, found in diachronic and ontogenetic studies of language development, in the creolization of Pidgins, as well as in SLA, there is a tightening up of syntax; constructions which were formerly organized to maximize pragmatic functions become grammaticized: word order becomes more rigid; there is an emergence of surface morphology such as agreement markers; and relative clauses develop out of TCs, conditional constructions, and interrogatives (cf. Haiman 1978). The ways and degree to which learners from particular languages accomplish and evidence this syntacticization appears, however, to depend on the extent to which their L1 was syntacticized, in Givon's (1979) sense of the word (L2 input considerations aside); and the effect of the differential syntacticization is reflected in the L2 forms used developmentally to achieve the L1 discourse functions.

1.4 Purpose of the Thesis: Statement of the Problem

The aim of this thesis is (1) to review some of the descriptions, explanations, and predictions that have been proposed to account for L1 discourse-syntactic transfer and the syntacticization of English IL, and (2) to conduct a data-based study to empirically determine the manner and extent to which language typology (i.e., Topic versus Subject Prominence) constrains the acquisition of ESL. The need to test Rutherford's (1983) model arises from several weaknesses this author has identified in the research methodology in Rutherford's pilot study, and also importantly, because no known research has attempted to replicate the study or to test the proposed hypotheses in a study with a rigorous design.

The three primary research questions addressed in this thesis are as follows:

6

(1)   What is the effect of the independent variables of Ll (Japanese or Mandarin) and proficiency level on the dependent variable of syntacticization of Topic in ESL?

(2)   To what extent does Rutherford's (1983) model capture and predict these effects?

(3)   Eow is the claimed transfer effect of Topic Prominence consistent with other constraints on language development?

1.5   Significance of this Research

The significance of this line of research is that it contributes to a better understanding, description, and account of processes in SLA and patterns that occur in J and M acquisition data. It assumes the importance of accurately conceptualizing, explaining, and predicting proposed IL developmental stages, and also advocates explicit and sound- research methodology in objectively testing the predictions. The research portion of this thesis is an attempt to improve on the methodology of the previous studies by explaining sampling and coding procedures# control variables, and analyses, as well as the interaction of the independent variables of Ll group and language proficiency with the selected dependent measures of syntacticization. Additionally, the study attempts to provide results which show more clearly than Rutherford (1983) has, the interrelationships and interaction among the different proposed stages of syntacticization. The possible role of expletives (dummy Subjects) and the acquisition of agreement as triggers in the shift from Topic to Subject Prominence will also be addressed. While Hilles (1985) and Rutherford (1983) have examined the expletive–trigger hypothesis, the function of agreement has not generally been examined adequately.

In short, this work pursues what many researchers currently consider to be a promising area of SLA investigation which considers the interaction of properties of universal grammar with language–typological and language-specific properties in the development of interim (IL) grammars and the production of a second language.

7

1.6. Organization of the Thesis

The thesis is organized as follows. Chapter II presents a literature review related to three fundamental sets of questions underlying this research: **(1)** What is transfer, what are its effects, and how is it constrained by other principles and processes of language acquisition? (2) What is a pragmatic and grammatical description of Topic Prominence and subject Prominence, and how are these constructs reflected in the grammars of English, J, and M? **(3)** How is the process of syntacticization characterized in studies of diachronic language change, and in L1 and L2 acquisition?

Chapter III presents the research methodology used to test Rutherford's model, and the analysis of the data. Chapter IV describes the results of the study, and Chapter V discusses the main findings. Lastly, in Chapter VI the answers to the research questions are 'summarized and conclusions are stated.

1.7 Limitations of this Work

In researching an area of SLA which has as many complex subcomponents as this one, it is impossible to furnish a complete treatment of each related issue. The thesis will, therefore, highlight what are considered to be the most important areas, and provide references to more comprehensive accounts of the phenomena in question wherever possible.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter is concerned with answering, in turn, the questions posed in Chapter I. These are: (1) How does transfer operate in SLA? (2) In what ways can English, J, and M be considered Topic and/or Subject Prominent? (3) How do natural languages generally evolve from Topic Prominence to Subject Prominence?

2.1 Transfer

In this section, we first introduce the concept of transfer, especially as it is used in terms of the constraints of Ll features or typological parameters in SLA. Second, we discuss how past studies have interpreted IL phenomena in terms of transfer effects or constraints. Third, we notice how transfer might be operative at the level of typological parametric variation? here in terms of discourse–level syntax.

We then examine several typological parameters which have been proposed in recent SLA studies as possible sources of Ll transfer in IL. We will summarize the role that Ll settings might play in the acquisition of L2 parameter settings, whether this is equivalent across the parameters or whether some are more powerful or have more prolonged effects in SLA than others. The typological parameters referred to are, roughly in order of intended emphasis: (1) Topic/Subject Prominence (T/Sp; Li and Thompson 1976); (2) Pragmatic/Grammatical Word Order (P/GWO; Thompson 1978); (3) PRO-drop (PD; Chomsky 1981, White 1984, 1985, Hilles 1985); and (4) Canonical Word Order (CWO; Greenberg 1966). Basically, all four share the common focus of the surface syntactic realization of Topics and Subjects; that is, they are not only concerned with the basicness of a function (e.g., Subject), but also its zero or overt surface coding, and its position in Ll and IL word order. Other typological parameters have been proposed in the transfer literature (e.g., Eckman et al. 1984, Gass and Selinker 1983), but for the purpose of the study in this thesis, those mentioned above which are concerned with Subjects, Topics, and word order in general are most relevant.

9

2.1.1 What is Transfer?

In the past work on transfer, much attention was paid to the "interference" of Ll surface syntax and morphology (e.g., in the production of the third-person "-s" morpheme in English, versus zero inflection), and the explanation was essentially a behavioristic one: patterns which were conditioned in Ll were also produced in L2 (see Gass 1984 for a review). However, more recent work has incorporated the theoretical frameworks of such fields as cognitive psychology, generative grammar, and developmental psycho-linguistics, and as a result, views the phenomenon in a more complex manner. Transfer effects are now considered to be less mechanical and more dynamic, integrated, and interactive with a number of other pervasive and powerful constraints on IL development.

Rutherford and Altman (1985:5) capture the difference in orientation of traditional and current discussions of contrastive analysis and transfer in the following way:

> studies have moved away from straightforward comparison of a second-language learner's IL and native language with respect to some strictly surface feature, and have instead begun investigating the possibility of more subtle influences of the Ll upon the shape of the IL . . . influences, for example, of Ll discourse function, of lexical features, of syntactic processes, of abstract organization, etc.

It is the purpose of the following sections to examine the current descriptions and explanations for Ll transfer# and, on the basis of a review of this literature, to account for why transfer of Topic Prominence is likely to occur in ESL.

Current accounts of language transfer define the phenomenon as

> a set of constraints [emphasis mine] that one's previous knowledge imposes on the domain from which to select hypotheses about the new data one is attending to (Schachter 1983:104).

Domain, as used here, refers to categories of abstract syntactic organization or knowledge, such as clause and phrase types, and lexical categories (p. 103).

10

However, transfer is not simply considered to be an independently operating process or set of constraints; as Zobl (1982) points out, transfer is itself constrained by a number of other interrelated factors and processes, including linguistic and psycholinguistic markedness (Rutherford 1982, 1984b), processing or parsing demands (Zobl 1982), (perceived) typological distance and the "borrowability" of the L1 with respect to the L2 (cf. Zobl 1985a, Kellerman 1983, 1984, Eckman 1981, Seliger 1984, Corder 1983, Jordens and Kellerman 1981), and core grammar (Rutherford 1984b, White 1984, 1985).

In the discussion which follows, a brief description will be given of some general effects of transfer in terms of overproduction, avoidance, and pace of IL development. Then, the factors listed above which are said to constrain transfer will be examined.

2.1.2 Transfer: Direct and Indirect Effects

Rutherford (1984b) states that transfer may eventuate in both direct and indirect effects in IL. Direct effects are generally reflected in the production of L2 structures with a more or less isomorphic correspondence with their L1 counterparts. Examples of direct transfer effects at the discourse-syntactic level are the literal translations of TCs into IL. In addition to cases of fairly transparent direct effects, however, there are many potential, often subtle, indirect effects. These can be described in terms of overproduction, avoidance, and pace of development.

Schachter (1983:104), cited above, views hypothesis testing as an integral part of transfer, although that position is not necessarily shared by others (cf. Gass and Selinker 1983). She predicts, in very general terms, that in second language learning

> Such phenomena as slower learning . . . , overproduction, and choice of wrong domain should be relatively more evident in the data of a learner of an unrelated target, whereas interference (choice of correct domain but wrong hypothesis) and positive transfer (choice of correct domain and correct hypothesis) should be more evident in the data of one who learns a related language.

Schachter's work has shown how learners may either overproduce structures for functional purposes, or simply avoid producing them altogether due to their apparent difficulty. An example Of overproduction was seen in Japanese and Mandarin learners' frequent production of IL *it* and *there*, which was unexpected since dummy Subjects do not exist in the Lls (Schachter and Rutherford 1979). The use of the term "overproduction" in this context does not necessarily mean that the learners produce more structures of a given type than native speakers would (although this could be tested with comparisons with normative data), but rather, that learners adopt certain forms to fulfill their intended (i.e., targetlike) function, and the same forms are overused or overgeneralized to fulfill other functions (e.g., topicalization) as well.

Also, in Mandarin and Japanese learners' ESL, the avoidance of relative clause formation (RCF) was considered by Schachter (1974) to be a consequence of the difference between RCF in those languages and English, which was apparently much greater than the difference between RCF in Arabic or Spanish and English. In the former case (i.e., M and J vs. English), the branching direction is different, whereas in the latter case (i.e., Arabic and Spanish vs. English), the problem is more restricted to the existence of resumptive relative pronouns in relative clauses (cf. Flynn 1984 for a discussion of the effects of principal branching direction in SLA). Several other researchers have also discussed the potential role and effects of avoidance in SLA (Kleinmann 1977, Hakuta 1976, Schachter and Hart 1979).

Zobl (1982) identifies the potential transfer effect as not only a matter of overproduction, avoidance, and unique types of errors in IL structures, but more significantly perhaps, that it is an SLA pace-setter of sorts. That is, the degree of contrast or similarity of L1 and L2 constructions determines the speed at which learners pass through certain inevitable developmental stages. We might extrapolate, therefore, that learners of both Tp and Sp languages will produce marked TCs in IL (see discussion below about TC production in L1 and L2 acquisition), but the Sp group will acquire the Subject-predicate structure of Sp English more quickly than the Tp group.

In Zobl's account, other transfer effects related to the pace-setting function are the initial point at which learners will begin to acquire a new construction (e.g., the particular stage in the four-stage developmental

sequence of English negation, as in Schumann 1979), and whether learners from an L1 which has constructions that are more or less congruent to those in the L2 can by-pass one or more of the earliest stages as a result. Furthermore, the persistence of some IL forms (i.e., the interruption of progress through developmental stages or fossilization at a non-terminal construction, e.g., no V or unanalyzed don't V, for negation) is another possible transfer effect. Thus, the research question posed by Zobl (1982:171) with regard to transfer is summarized as follows:

> whether different language groups, faced with a particular structural domain of the L2, begin the developmental continuum at the same point, move through it at the same pace, and traverse the same developmental structures in the same sequence.

In terms of discourse-level syntactic transfer related to Topic Prominence, it is possible in the perspective outlined above to predict that L1 speakers of J, which has a more syntacticized Subject than M, will produce more grammatical Subjects in early stages of ESL. This is because their L1 typology might enable them to proceed toward Subject Prominence at the initial developmental point of NP-NP-VP (i.e., Topic-Subject-Predicate), as opposed to, say, M, who might proceed from an initial stage of NP-Ø-VP (i.e., Topic-Predicate) according to Rutherford's six-stage sequence described later in this chapter.

2.1.3 Transfer, Natural Acquisitional Processes, and Perceptual Operating Principles

Andersen (1983:177) states that transfer of L1 structures serves to filter learners' perception and retention of L2 input. Following up on some of Zobl's work, he conceives of transfer as the interaction of a number of perceptual operating principles (cf. Slobin 1977), L1/L2 structural differences, and natural acquisitional principles. An example which Andersen (1983:181) cites to illustrate this point follows. In the development of negation by speakers whose L1 has preverbal negation (i.e., no V), two forces combine to "promote" IL no V: (1) natural developmental sequences (i.e., stage one negation in English is no V), and (2) transfer of the L1 negation scheme. By contrast, for learners whose L1s have postverbal negation, preverbal negation in IL is "less frequent and less enduring."

13

In addition to the role of L1 structure, Andersen (1983:182) acknowledges the input structure and the relative frequency of the structure in the input to be a factor in transfer, especially when there already exists within the L2 input the potential for (mis-)generalization from the input to produce the same form or structure" (as in natural developmental sequences). For example, Andersen could account for the early acquisition and overgeneralization of the English locative preposition in for a Spanish speaker (Marta), and her late acquisition of the possessive "-s" morpheme, as a result of the relative frequency of the two structures in English input, and the degree of similarity of the L2 structures with their L1 counterparts. That is, in is very productive in English, compared with, say, on, and it resembles the Spanish preposition en; however, "-s" is relatively unproductive, and does not correspond as closely to the Spanish possessive.

Within Andersen's analytical framework, we might posit that because TCs occur in English as marked constructions (especially in colloquial English), this promotes their occurrence in IL. Furthermore, the sometimes shared properties of definite, sentence-initial Subjects and Topics might also cause learners to produce marked Topics in ESL. Another feature of Modern English is that agreement marking is limited to third-person "-s" only, and this might account for the fact that the acquisition of agreement is generally rather late as compared with other morphemes. Also, it is true that some constructions considered ungrammatical in Standard English (e.g., Rutherford's stage (iii), 'There are a lot of people get married after 40") occur frequently, nonetheless, in some dialects. Lastly, the Foreigner Talk to which learners are often exposed also contains evidence which supports the (misguided) notion that English is Topic Prominent.

2.1.4 Transfer and Markedness

Rutherford (1982) discusses syntactic, psycholinguistic, and typological markedness as important elements in transfer. He claims that learners will attempt to strike a balance between complexity and explicitness of structures in IL, based on markedness principles; thus unmarked forms tend to be acquired prior to marked ones on a developmental continuum for any given structure. Learners will also attempt to make structures less complex by rendering them more transparent, by means of added redundancy (where it would otherwise be optional or perhaps ungrammatical--e.g., resumptive pronouns in relative

14

clauses, left dislocations, or other kinds of Topic reinstatements, as in Chaudron 1983), and by reducing syntactic complexity through coordination as opposed to subordination.

In this light, we might account for the appearance of pragmatic discourse-functional strategies in IL as a sign that learners are applying principles of transparency, explicitness, and non-complexity.

Kellerman (1983) posits two related perceptual strategies employed by learners in determining whether a particular structure can be effectively transferred. Both strategies derive from what Kellerman terms "psychotypology[n]"; that is to say, the learners' perception of the typological distance between L1 and L2. First, learners unconsciously measure the apparent distance between the L1 and L2. Second, they must ascertain the relative markedness of the L1 structure, that is, how "reasonable[n]" an entity it is for the purpose of positive transfer. Implicit in this judgment process is the notion of whether a certain L1 form is perceived by learners as language-specific, and thus non-transferable, or conversely, language-neutral, and thus transferable.

First Kellerman illustrates the psychotypological hypothesis with data from Finnish and Swedish speakers learning ESL. Finnish is very different linguistically from either Swedish or English and, according to Kellerman, Finns perceive this typological distance and thus refrain from transferring various Finnish-specific structures. Swedes, however, are more apt to transfer structures (sometimes erroneously) from their L1 because Swedish and English, the target language, are considered typologically close.

Second, examples related not only to the perception of typological distance, but also to the sensitivity to relative structural markedness are discussed in the literature. In some cases, psychotypological principles are found to operate in terms of semantic properties of verbs and idioms, and whether the L1 usage will be considered a "reasonable entity" to be used in IL (Jordens 1983, Kellerman 1983, Hawkins 1980). Kellerman (1983:119-120) cites one such example, whereby

Dutch learners of English will reject sentences in English of the kind John is easy to convince although they are "modelled" by the perfectly normal Dutch Jan is makkelijk te overtuiaen, on

15

the grounds that "it is not Jan who is easy: it is convincing him that is."

Similarly, Gass (1984:124) emphasizes that transfer is not a "bidirectional," symmetrical process, such that structures are mutually transferred from one language to another as mirror-images. Citing examples from Swain et al. (1972), Gass (1984) pointed out that while English learners of French produce erroneous constructions with a final clitic pronoun, as in (a) below, French learners (whose L1 places object clitic pronouns preverbally) do not produce constructions with preverbal object clitics in ESL, as in (b):

a. * Le chien a mangé les (produced).
   the dog has eaten them

b. * The dog them ate (not produced).

In other words, the French structure, as exemplified in (b), is more marked than its counterpart in (a), therefore it is not transferred. This illustrates again that typological distance is one potential criterion for transfer, but that it is closely constrained by markedness principles. Since the L1-L2 typological distance is equivalent between French and English, the perception of the relative markedness of the placement of object clitics seems to be a more powerful constraint; thus distance may, in some cases, be a necessary condition, but it is not always a sufficient condition for the occurrence of transfer.

Kellerman (1983:122) describes the tendency toward linguistic transparency and non-complexity in the following way: "In the absence of specific knowledge about the L2, learners will strive to maximize the systematic, the explicit and the "logical" in their IL."

The logical side of transfer is illustrated in the discussion of IL conditional constructions, whereby second language learners tend to produce strings such as (a) below, rather than (b). Would have is considered a logical choice for the verb phrase in the first clause in (a) because it refers to a counterfactual event rather than a real event which has occurred in the past.

a. If I would have (VP), I would (VP) . . .

b. If I had (VP), I would (VP) . . .

16

The same principles were said to be operative when learners were asked to translate into English the Dutch equivalent of the sentence 'His fall was broken by a **tree**," which represents figurative or idiomatic speech in the **two** languages. Evidently, learners perceive the 'mismatch[n] between the literal and figurative meanings of idiomatic expressions[r] and **thus**, their relative markedness **(p. 123)**; therefore[r] they produced such sentences as **"His** downfall was eased by a **tree**," reflecting an **"attempt** to reduce metaphorical opacity in the L2 by an avoidance of structural correspondence between Dutch and English[n] **(p. 124).**

Gass **(1979)** also noted the tendency for learners to transfer logical constructions such as resumptive pronouns in relative clauses[l] which are present in English deep or logical structure[l] but not in surface structure.

This psychotypological perspective on transfer would lead us to believe that learners of Tp languages feel free to carry over **TCs** into **ESL,** not because they perceive the languages to be typologically **close,** but rather[l] because it is pragmatically natural and logical to present known information before unknown[r] or to frame one's utterance before making assertions. **Kellerman** (1983) observes that there is even a kind of iconic logic to this pattern: what is uttered earliest is already 'older[n] and "more given" than what follows it serially in an utterance or in discourse. In American Sign Language (Bates and **MacWhinney 1979),** and in modalities of communication other than language, this pragmatic strategy is also apparently utilized; **e.g.,** in dance[l] mime[r] and advertising **(Mallinson** and Blake **1981),** and in 'canine[n] communication **(Givon 1979).**

Indeed[r] citing the Schachter and Rutherford (1979) finding that properties of L1 Tp are transferred in **ESL,** Rutherford **(1982:104)** speculates that[r] in spite of the fairly obvious typological distance between the Tp and Sp languages[r]

It might plausibly be argued **. . .** that the discourse features in question (**viz.** those having to do with the basicness of topic) are unmarked[l] and that in such a situation the markedness parameter **prevails,** and therefore transfer occurs.

2.1.5 Transfer and Typological Parameters

In the foregoing section, we observed how perception of typological distance constrains transfer, even in cases where both L1 and L2 have equivalent, marked expressions such as in the Dutch and English expression "His fall was broken by a tree." Furthermore, it was suggested that learners might transfer L1-like TCs into English, due to their functional transparency and presumed aid to processing. Below, transfer is considered in terms of several other typological distinctions, with reference to word order, in particular.

With regard to typological transfer, and especially word order-related typologies, many applied linguists have studied more global transfer effects, at the organizational level of word order. Languages have been typed variously, according to canonical word order (CWO), i.e., the serial arrangement of Subject (S), Verb (V), and Object (O) (Greenberg 1966), and whether word order is used to convey pragmatic (PWO) or grammatical (GWO) information (Thompson 1978).

Studies have shown that the more purely "syntactic" word order parameter or arrangement is unlikely to be transferred from the L1, at all but the most basic levels. That is, the L1 serial order of the constituents S, V, and O is rarely carried over into IL; on the contrary, there, appears to be little difficulty for learners to master word order and adposition order at low levels (Curran 1984, Hawkins 1980, Lehmann 1973, Rutherford and Altman 1985, Lujan, Minaya, and Sankoff 1984). Presumably, there is a great deal of (comprehensible) input to learners that attests to the order of S and O relative to V, and also closely connected to this, the order of adpositions (prepositions, postpositions) with respect to head NPs. In pidginization data there seems to be a greater frequency of deviant canonical word order in "basilangs" (Schumann, forthcoming), which can be attributed to relexification of the L1, but this is still relatively uncommon in SLA on the whole.

To empirically test this effect of the non-transferability of CWO, Rutherford and Altman (1985:5) studied ESL written production data of a total of three hundred Japanese (SOV), Arabic (VSO), and Spanish (SVO) learners. They found that it was not CWO but rather the effects of L1 PWO and Tp that were most evident in IL. They conclude, therefore, (p. 5) that

L1 (syntactic) canonical word order does not undergo transfer in the L2 learning experience, and it is implied that this apparent transfer power of discourse as opposed to syntax is not coincidental.

Zobl (1985b) posits a naturalness order for the acquisition of the three typological parameters in ESL: CWO (SVO) > PRO-Drop (-PD) > Tp/Sp (Sp). Hence, L2 CWO is acquired first, followed by the correct setting of the PRO-drop parameter (-PD for English), and finally, Subject Prominence is acquired. It could be inferred from Rutherford and Altman's (1985) results that a parallel typological shift, from PWO to GWO, occurs somewhere between the acquisition of the correct PD setting and Sp; and, like the Tp parameter, the PWO parameter has a prolonged presence in IL.

Besides Rutherford and Altman (1985) and Zobl (1984, 1985b), other linguists working within the theoretical framework of Government and Binding (GB) have studied the setting of the PRO-drop (PD) parameter in L1 and its possible transfer effects in IL. For example, White (1984, 1985) and Hilles (1985) investigated how an L1 setting of +PD bears on the acquisition of the correct L2 setting in SLA. Furthermore, these researchers point out that one of the reasons that typological parameters established in L1 are such powerful sources of transfer is that (1) the parameters are part of Universal Grammar (UG), and once the L1 setting is fixed one way based on "positive evidence" or language input to the learner, a considerable amount of counter-evidence might be required to reset the parameter in SLA; and (2) each typological parameter theoretically comprises a number of constituent properties; e.g., PD not only involves Subject PRO deletion, but also (hypothetically at least) that-trace violations and Subject-verb inversion as well.

what is of interest here is how these researchers account for the influence of the L1 setting in IL, in terms of the "positive evidence" that is required in order for learners to acquire a target language configuration or parameter setting (i.e., to "reset" the existing parameter); implicit in this notion, of course, is the theoretical and experimental interest in *parametric variation" that is said to exist and play a role in the SLA process (cf. White 1985).

This approach, as espoused by White (1985:48), is summarized as follows:

19

Whilst it is not necessarily the case that second language (L2) learners approach language learning in the same way as L1, they are faced with an essentially similar task, namely that of inducing a grammar on the basis of data which are impoverished in various ways. Thus, it is worth investigating whether UG plays a role in L2 acquisition and, if so, whether the L1 affects the way that UG operates in L2. The concept of parametric variation is of particular interest where L2 acquisition is concerned, since L2 learners will often be in the situation where their L1 has fixed some parameter one way, whilst the target language has some other setting, or the situation may arise where the first language has some parameter activated which is not operative in L2, or vice versa.

White (1984, 1985) conducted studies to measure the influence of the variation of the PRO-drop (PD) parameter in SLA. There are two possible settings which languages can have: +PD, (i.e., null-Subject is a common feature in the language), or -PD, (i.e., Subjects must usually appear in surface syntax). In both studies, White reported that a group of Spanish learners with the L1 parameter setting of +PD produced or judged acceptable more instances of Subject pronoun deletion in ESL than a control group of French speakers, whose L1 is -PD like English. Other PD properties such as that-trace violations (i.e., the extraction of Subjects out of clauses containing a complementizer) and free Subject-verb inversion were also examined. In short, the difference in the PD setting in the L1 and L2 (i.e., the parametric variation) led to judgments which demonstrated that the L1 parameter was still operative in IL.

Hilles (1985), working within the same linguistic framework as White, drew attention to the role of possible "triggers" or pivotal points in the development of ILs in terms of the (re)setting of typological parameters. She found that the expletive there was an example of one such trigger in the transition from +PD to -PD in a Spanish adolescent's ESL. Hilles' data also revealed that there was an inverse developmental relationship between frequencies of +PD and modals.

Lastly concerning the PD parameter, Gundel (1980) and Gundel and Tarone (1983) suggested the interaction of a Topic Prominent parameter setting with the occurrence of PD. This point will be relevant in the discussion of our

20

research findings in Chapter V, because we can account for Topic chains and null-Subject (PD) as IL reflexes of **both** L1 Tp and L1 PD.

In summary, on the basis of studies looking at typological parametric variation between L1 and L2, such as those cited above, it can be seen that the (re)construction of an IL grammar at the macro-Level of typology is in some cases a straight-forward, immediate process (e.g., CWO), accompanied by little deviation from the L2 norm. In other cases, however, (e.g., PWO, PD, Sp) and especially at the discourse-syntactic level, it is apparently a much more gradual, demanding process, perhaps due to computational or parsing demands. Another possible explanation, which was not explicitly discussed above except in relation to the cited implicational sequence of CWO > PD > Sp, is that there are a number necessary grammatical prerequisites in the process of syntacticizing Topics in SLA (cf. Zobl 1985b).

2.1.6 Transfer and Universal Grammar

Closely related to the discussion of language typology, those linguists familiar with language universals have postulated and thus far found support for the notion that regardless of **why** transfer occurs (the **particular factors involved**), its **effects** in terms of emergent IL structures will not violate **principles and sequences** documented for natural **languages--in** diachronic, ontogenetic, phylogentic, **and synchronic** linguistic **data** (cf. Lujan et al. 1984, Hawkins 1980, Zobl 1984, Schmidt 1980, Gass and Ard 1984, Eckman et al. 1984). These universal, natural developmental and synchronic tendencies will, therefore, be described for the Tp-to-Sp shift in the third part of this chapter.

In conclusion, in the present research which examines typological constraints on the process of SLA, in order to meaningfully gauge the strength and direction of the potential typology-based transfer effect, we need to consider four general research objectives. First, we must determine under what circumstances or in what environments learners will apply the typological parameter settings of their L1 in IL. Second, we must determine how the application of parametric constraints manifests itself in IL, for example, by looking at cross-sectional studies of learners[1] progress at various stages of development (i.e., proficiency), which might also show a shift from one parametric paradigm to another.

Third, attempts should be made to explain why some features have constraining effects in IL while others do not; that is, why CWO apparently is an insignificant source of transfer[r] compared to PWO and Tp, or why Tp should be so persistent in ESL data. As we have seen, this issue is addressed in the work of applied linguists cited above, who have incorporated in their discussions principles of markedness theory, language universals, historical linguistics, and language acquisition data. As reported[r] they observe the emergence of common patterns of natural developmental sequences in data from a number of different diachronic and synchronic sources.

Last, it is important for the description and explanation of the phenomenon to have predictive power (cf. Rutherford 1984b, Gass 1984), such that the findings can be replicated and generalizations can be made. Again, an important reason for testing Rutherford's model in Chapter III is to determine whether it has predictive power and is, thus, generalizable.

2.1.7 Summary of Part I

Why would learners transfer features of a Tp L1 to an Sp L2 as hypothesized by Rutherford (1983), and what would the effects be? Below are listed a number of plausible reasons which have been drawn from the review of transfer presented in the foregoing discussion:

(1) Overproduction: Learners overproduce certain IL structures because they are used to achieve discourse-functional strategies in the Ll; or they overproduce forms which their Ll typology apparently sensitizes them to, as in J production of extraposition.

(2) Avoidance: Learners avoid the production of Subject-verb agreement, dummy Subjects, and relative clauses (i) due to the lack of equivalent or comparable structures in their Lls, and (ii) due to Ll Tp, PWO, and PD, which involve the deletion of Subjects (as in Tp, PD), Subject-verb inversion (PWO,PD) and lack of agreement (Tp).

(3) pace-setting: There is a prolonged effect of Tp due to the consistency of features of Tp (e.g., TCs) with certain psychological operating principles (e.g., to make topical information salient and place it in sentence-initial position) and natural acquisitional processes, along with the basicness of the pragmatic mode in communication; furthermore, the occurrence of TCs in

colloquial English in particular may provide conflicting
evidence for the learner as to the normative usage and
basicness of Sp.

(4) <u>Markedness</u>: TCs are transparent and possibly
unmarked (but see Givon 1984); in spite of typological
distance, learners do not perceive TCs to be opaque or
complex, and this is perhaps also due to their pragmatic
effectiveness.

(5) <u>Word **Order** **Typology** and Parametric Variation</u>: Tp
is the last of a sequence of grammatical and functional
word order–related typological parameters to undergo a
resetting to conform to English; the shift from Tp to Sp
must be preceded by the acquisition of English CWO (SVO)
and PD (-PD).

(6) <u>Universal Grammar</u>: As mentioned above in point
(2), Tp and the stages toward Sp are consonant with
universal developmental sequences, which will be discussed
later in this chapter.

2.2 Topic Prominence and Subject Prominence

Central to this thesis is the assumption that some languages operate in a more "pragmatic" mode than others, that is, they are more "discourse-oriented." This is reflected in the basic sentence structure of the languages, and in terms of the role of Topics in particular. In this section, we outline proposed linguistic differences between Topics and Subjects, and Topic Prominence (Tp) and Subject Prominence (Sp), by first presenting Li and Thompson's (1976) typological classification. Secondly, we review other literature in which an attempt is made to operationalize these terms. Thirdly, we describe TCs in English, J, and M.

In contrasting the three languages, we are interested in highlighting the degree to which TCs are basic to J and M, and thus unmarked constructions, compared with their counterparts in English, which are considered relatively marked. Furthermore, the kinds of categories which will be tested for in the research portion of the thesis will be briefly examined here. We will provide some explanation as to why locative phrases and conditionals in J-ESL are said to reflect Topicalization strategies from the L1, and why Topics are claimed to be so abundant and productive in M-ESL. We will also present examples of marked and unmarked Topic-Comment constructions in English, predicting that learners will produce marked TCs before they learn to consistently produce less marked ones.

2.2.1 A Typological Description of Tp and Sp

Li and Thompson (1976) characterize languages according to the basicness of the relation of Topic (T) and Comment (C), on the one hand, or Subject (S) and Predicate (P), on the other, conceding that some languages share features of both Tp and Sp or, conversely, neither Tp nor Sp. Exemplifying the various combinations are: Mandarin Chinese as Tp; English as Sp; Japanese as Tp and Sp (TSp); and such Philippine languages as Tagalog and Ilocano as -Tp and -Sp (-TSp) (see Figure 2.1).

24

Figure 2.1

Topic/Subject Prominence:

Li and Thompson's (1976) Typology

(A) Tp

Topic notion integrated
into basic sentence structure;
Topic and Subject distinct

(D) Both Tp and Sp

Topic sentences become
less marked, more basic

(B)   Neither Tp nor Sp

Topic   becomes more
closely integrated into
case frame of verb

(C) Sp

Topic has become integrated
into case frame of verb as a
Subject; Subject and Topic
often indistinct, Subjects
having some non-Topic properties;
sentences with clear Topics are
highly marked

Li and Thompson operationalize the above distinction by pointing out that the prototypical Topic (1) is a definite (or generic) NP; (2) is underived from another constituent NP (i.e., it is basic); (3) has an optional selectional relation with the verb; (4) has a functional role of specifying the domain of the predicate; (5) does not govern agreement with the verb; (6) receives sentence-initial surface coding; (7) does not have grammatical control over such operations as EQUI-NP deletion (EQUI), reflexivization (REFL), or imperatives (IMP); and (8) controls coreferential constituent deletion.

Other characteristics of languages with Tp are, according to Li and Thompson (1976), that they (9) only rarely have passives, or there is a very narrow range of passivization (e.g., adversive meaning); (10) have no dummy Subjects; (11) allow "double Subject" constructions; (12) are typically verb-final languages.

In contrast to the functional characteristics of Topic, according to Keenan (1976), Subjects (1) are either definite or indefinite; (2) have selectional relations with verbs; (3) have a primarily semantic role; (4) control verb agreement; (5) may occur in other than sentence-initial position; (6) govern numerous grammatical operations in predicates, e.g., EQUI, REFL, IMP. Finally, unlike Tp languages, Sp languages typically have (7) passives and (8) dummy Subjects.

It is not entirely clear, however, to what extent TSp or -TSp languages differ with regard to each of the above functional features, although Li and Thompson (1976) present them along a continuum that is reproduced in Figure 2.1. TSp languages have basic sentences in which both Topic and Subject figure prominently, on the one hand, and conversely, -TSp languages have basic sentences in which neither Topic nor Subject are primary in the sense that they have been described above. Note, for example, that in addition to having a grammatical category of Topic, J has Subjects coded with the postposition ga, and has a wider range of passives than pure Tp languages, but that Subject in Japanese does not trigger agreement marking on the verb; notice also that Philippine languages have a Topic/Focus constituent in phrase structure, but that it does not assume all of the functions or criteria associated with Topic (e.g., sentence-initial position) in Li and Thompson's framework (cf. Foley and Van Valin 1984).

2.2.2 Topic-Comment Constructions

In this section we outline attempts in the literature to operationalize the TC construct which, we predict, is evidenced in J- and M-ESL due to both universal (e.g., pragmatic) constraints, and L1 typological constraints. It is, however, beyond the scope of this typological sketch to present an elaborate treatment of issues related to the identification or derivation of TCs or to provide an exhaustive contrastive description of English, J, and M along the same lines. For a more complete treatment of topicalization in these languages, see, for example, Gundel (1977), Magretta (1977), Reinhart (1982); Duff (1980), Kuno (1973), Kitagawa (1982), Farmer (1984); Barry (1975), Xu and Langendoen (1985), and Li and Thompson (1981).

We will adopt Li and Thompson's (1976) grammatical criteria presented above to distinguish Tp and Sp languages, and to identify Topics and Subjects.

Syntactically, Topic is usually conceived of as being set apart from its comment as a sister node to S, as the following phrase structure rules indicate (Chomsky 1977:91):

$$S^n --> \quad TOP \ S'$$
$$S' --> \quad COMP \ S$$

Xu and Langendoen (1985:17) suggest, however, that for Chinese the appropriate rule is

$$S' --> TOP \ \{S, \ S'\}$$

In several current syntactic descriptions, the Topic node is considered to be basic in Mandarin and Japanese, and the Topic is bound to its predicate by means of a "Topic binding" operation. Where there are coreferential elements in the predicate, coindexing occurs on a pronominal (PRO) element; in cases where there is no coreferential element in the predicate, Topic binding is achieved pragmatically (Kitagawa 1982, Farmer 1984).

In English, on the other hand, Topic is generally considered to be a constituent that is preposed by a leftward movement rule, as in left dislocation, as the following rules from Ross (1967) illustrate:

27

<u>Topicalization</u>   <u>Left</u> Dislocation

```
    X  NP  Y                        X  NP  Y
    1  2   3  ==>                    1  2   3  ==>
 2  #  1   0  3                   2  #  1   2  3
                                          [+pro]
```

Since adverbial preposing is achieved through an
operation very similar to Topic binding or topicalization,
preposed prepositional or postpositional phrases and
adverbials are considered to be special cases of
topicalization (Xu and Langendoen 1985, Chafe 1976, Barry
1975, **Li** and Thompson **1981).**

The reason these sentence-initial time and
locative phrases are considered to be topics is
simply that they have all the properties
of topics: they set the frame within which the
rest of the sentence is presented, they are
definite, referring to places, times about which
the hearer already **knows,** and they may be followed
by a pause **(Li** and Thompson **1981:95).**

Ross **(1967)** represents the adverb **preposing**
transformation as **follows,** although again, some discussions
of Mandarin and Japanese **posit** an underlying (basic)
adverbial Topic in deep structure.

Adverb <u>Preposing</u>

```
      X - +adverb - Y
      1      2      3  ==>
   2 + 1     0      3
```

Not only sentence-initial adverbials and **NPs** appear to
serve as Topic, though. As Haiman (1978) points out,
clauses such as conditionals have a similar pragmatic
function; and in some languages, in their diachronic
derivation and synchronic coding they are linked to Topics
and Interrogatives. Haiman **(1978:583-585)** thus considers
conditionals to be a special kind of Topic construction:

A conditional clause is (perhaps hypothetically) a
part of the knowledge shared by the speaker and
his listener. As such, it constitutes the
framework which has been selected for the
following discourse.

Lastly, it is generally accepted that a sentence may
have more than one Topic, since individual clauses within

28

the sentence may have their own topics; thus, there is sometimes a "nesting" or "layering" of Topics, especially in complex sentences (Mallinson and Blake 1981, Xu and Langendoen 1985, Kitagawa 1982, Kodama 1981).

Above, some of the syntactic properties of Topics were presented, and also, the potential for adverbials and conditional phrases to appear sentence-initially as Topics or topicalizers. It was suggested that Topic has certain pragmatic functions, and here, this point will be considered further. A number of primarily functional linguists have characterized Topics and Comments variously, according to a number of bipolar terms (cf. Bates and MacWhinney 1979), such as given-new, presupposed-asserted, and activated-unactivated (Gundel 1978), which are semantic and pragmatic in nature. Three commonly cited criteria of Topics are (1) the notion of "givenness"; (2) the principle of "aboutness" expounded by Reinhart (1982) and Gundel (1977); and (3) the "framing" principle (cf. Chafe 1976), already alluded to in the discussion of conditionals.

Gundel (1978:2) describes Topic in terms of givenness, as follows:

> The Topic, is 'given' in that it represents the starting point of the sentence, what the sentence is about. It is also given in the sense that its existence or truth is something that is already assumed to be known; i.e., it must be part of the speaker's and addressee's general knowledge if something is to be felicitously communicated about it.

Aboutness refers to the notion that the Topic is that element about which the Comment or assertion is made. The test for aboutness is whether the element in question can occupy an "as for X" or an "about X" clause, and still preserve the meaning of the sentence. This factor of 'aboutness' has recently been formalized by Gundel (1985:1), on the basis of a large cross-linguistic survey, as follows:

> An entity, E, is the topic of a sentence, S, iff in using S the speaker intends to increase the addressee's knowledge about, request information about, or otherwise get the addressee to act with respect to E.

Thirdly[8] Topics are said to be framing or scene-setting elements, such that they

> are not so much "what the sentence is about" as "the frame within which the sentence holds" (Chafe 1976:51).

Topic, thus identified, "sets a spatial, temporal, or individual framework within which the main predication holds" (Chafe 1976:50).

Lastly, Barry (1975:3) describes the framing function of Topics as follows:

> the grammatical target slot topic has the function of specifying the selevant universe of discourse (frame of reference, domain of referentiality) of its comment; the range of things with respect to which it makes sense to assert the comment.

By some accounts of what constitutes a Topic,' "aboutness" and "framing" are considered mutually exclusive features, and Topichood is borne by one or the other but not by both. Other accounts accept both criteria as qualities of Topics, and argue that there are simply different kinds of Topics. The latter view will be assumed in the present discussion, since structurally and pragmatically the two kinds of Topics are roughly comparable.

In the discussion which follows, English, J, and M will be described in terms of their typological and syntactic characteristics.

2.2.3 English

English is an SVO language, in which word order is relatively inflexible. For this reason Thompson (1978) refers to it as a language with Grammatical Word Order (GWO), in which structure preserving rules operate to ensure that the preverbal Subject position is always filled; these rules include passivization, raising, Tough movement, and focus constructions, such as cleft and pseudo-cleft sentences. In Li and Thompson's (1976) typology, English is cited as a typical Sp language. Topics in English tend to be "grammaticalized," and those which occur external to the sentence arise thsough application of a movement rule.

Some Prague linguists contrast Theme and Topic in that Theme specifies the universe of discourse (the frame), whereas Topic "presents the entity *about' which the predication predicates something in a given setting" (de Groot 1981:75). For example, in the sentences,

a. That trunk, put it in the car

b. As for Paris, the Eiffel Tower is really spectacular

in de Groot's view, the underlined elements are Themes, and it and the Eiffel Tower are Topics, in (a) and (b), respectively. Indeed, in English the semantic Topic is generally identical to the syntactic Subject; for this reason Li and Thompson (1976:484) explain that

subjects are essentially grammaticalized topics; in the process of being integrated into the case frame of the verb (at which point we call them subjects), topics become somewhat impure, and certain of their topic properties are weakened, but their topic-ness is still recognizable.

Consequently, in English we are able to identify both weak (i.e., diachronically integrated) Topics, which are unmarked, and strong Topics, which are the result of topicalization. In sentences (a) and (b) above, therefore, the underlined sentence-initial elements (de Groot's "Themes"), are topicalized constituents, filling the marked Topic position; the sentence-internal Subjects (de Groot's "Topics") are in the unmarked pre-verbal position. Other examples of the strong-weak Topic distinction will be presented below.

In many theoretical accounts, English TCs have been differentiated on the basis of marked and unmarked Topics/Themes (the terms will be used interchangeably hereafter), and in examining IL production of English constructions and the syntacticization of Topic, this distinction is meaningful. For at the earliest stages in Rutherford's (1983) model, learners typically produce "marked" Topics, and only at the final stages do they consistently produce the "unmarked," targetlike, integrated Topic/Subject.

Halliday (cited in Kress 1976) exemplifies the marked/unmarked distinction with the following sentences:

31

Marked Topics

    a.   <u>These houses</u>, my grandfather sold
    b.   <u>In</u> <u>Madrid</u> **it** was terribly cold

Unmarked Topics

    a.   <u>The</u> <u>catalogue</u> had obviously been prepared
        before they'd hung the pictures
    b.   <u>John</u> is the leader

Other examples of marked English TCs from Magretta (1977:55) [see also note 2], Gundel (1977:133), and Thompson 1978:32) follow; usually the preposed element is separated from the main sentence with juncture, such as a pause, and by appropriate changes in stress and intonation. Also, the Topics are often introduced with such phrases as 'Concerning X," "As for X," "You know X?, He/she . . .".

<u>Topicalization</u>

    a.   Scrambled eggs **I** can't stand to look at in the
        morning
    b.   The window John broke
    c.   John **I** didn't expect to run into
    d.   **A** kangaroo Jim **claimed** he has never **seen**
    e.   Cigarettes **I** don't think **I'll** ever be able to
        give up
    f.   That paper on topicalization Max told Sheila not
        to tell anyone he had written

<u>PP/Adverb</u> <u>Preposing</u>

    a.   Into the room flitted Mrs. Goldstein
    b.   Yesterday Mary went swimming
    c.   In the park we saw some jugglers
    d.   About his wife we know little of interest

<u>Left</u> <u>dislocation</u>/ **"**<u>As</u> <u>for</u>**"** Constructions

    a.   John, **I** know him
    b.   Ronald Reagan, **I** wonder if the voters really
        like him
    c.   As for tobacco, the government may raise
        taxes on **it**
    d.   About the tax loopholes, most citizens are
        unaware of them

The last category[1] left dislocation (which, in Magretta 1977 includes a special class of constructions with 'As for", as in (c) and (d)) is, therefore[1] another form of topicalization. However[1] it is distinguished from other kinds of topicalization in that there remains in the main sentence an overt pronoun that is coreferential with the preposed NP.

2.2.4 Japanese

Compared to English, which is Sp, and M, which is Tp, Japanese is said to have base structures which are both Tp and Sp, suggesting that Subject is more fundamental to J sentence Structure than it is in M. Although Topics in J and M share many common features[1] in J (but not M) Topics are morphologically coded: the postposition wa marks Topics; Subjects are also morphologically coded in J, with the postposition ga.

According to Thompson (1978), J is both a PWO and GWO language[1] whose CWO is SOV. Word order is flexible to some extent due to the elaborate case marking system, although changes in word order position express different pragmatic/semantic interpretation. Because J is a rigid verb-final language[1] however, this word order flexibility is somewhat constrained. Recall that Rutherford (1983) attributed some of the observed differences in J and M ESL data to this partial GWO status of J, namely that J were more sensitive to filling the Subject position in English with dummy Subjects.

In addition to the surface coding of Subjects in J, and the nature of J word order, a third feature which differentiates it from pure Tp languages (like Mandarin or Lisu) is the variety and frequency of passive constructions that occur in J. There are both 'indirect" and "direct" passives[1] with either adversive or neutral semantic features[1] depending on the construction (Kuno 1973; Howard and Niyekawa-Howard 1976). Furthermore, J demonstrates Sp features of Subjects with selectional restrictions and grammatical operations within the sentence, such as EQUI, naaara-constructions[1] reflexivization, honorification, and quantifier float (Duff 1980).

Some of the features most commonly associated with Tp and which occur in J are: surface coding of Topic; 'double Subject' constructions (two consecutive nominals, NP1-wa/ga, NP2-ga); verb-final CWO; no dummy Subjects; and no syntactic Subject-verb agreement coded on the verb.

33

The sentences below illustrate some of the features related to the TSp classification of J. In (a), there is a typical example of a "double Subject" construction (as in Li and Thompson 1976); the first "Subject" is topicalized and thus coded with wa, perhaps to avoid the ambiguity of two ga-marked "Subjects" (Martin 1975), but the whole–part relationship between the two nominals remains. In (b), a topicalized locative appears sentence-initially without a locative case marker. In (c) the lack of expletives in meteorological-type constructions is illustrated; whereas in the English gloss the dummy pronoun it is used, in J the Subject NP is ame 'rain'.

a. Sakana–wa tai–ga oisi–i
   fish–TOP red Snapper–NOM delicious–PRES
   '(As for) Fish (Topic), red snapper is delicious'

b. Gakkoo–wa boku–ga isogasi–kat–ta
   school–TOP I–NOM busy–PST
   '(As for) School (Topic), I was busy'

c. Ame–ga futte–iru
   rain–NOM fall–PROG
   'It's raining'

Kitagawa (1982) presents a number of other examples of TCs in J; among those cited are the following (p. 176):

a. Naomi–wa udon–o tabe–ta
          TOP noodle–ACC eat–PST
   'Naomi ate udon noodles'

b. America–de–wa Sumiko–ga kuruma–o kat–ta
          at–TOP   NOM car–ACC buy–PST
   'In America, Sumiko bought a car'[1]

c. Taroo–wa Hanako–ga iede–shi–ta
          TOP   NOM leave–home–do–PST
   'As for Taro, Hanako ran away from home'[1]

d. Bunmeikoko–wa   dansei–no heikin zyumyoo–ga
   civilized–nation–TOP man–'s average life span–NOM

       naga–i
       long–PRES

   'Civilized countries, their male population's
    average life span is long'

Again, as above, Topics may be the first part of TCs resembling "double Subjects" (c,d); whereas the whole-part relationship is quite clear in (d), it requires more pragmatic interpretation in (c). Alternately, the TCs may involve preposed adverbials and postpositional phrases, as in (b), or topicalized Subjects or other constituents, as in (a).

Within the X-bar syntax convention, Kitagawa (1982) posits a pragmatic "Topic Binding" formula or operation, by means of which Topic phrases are linked with coindexed elements (PRO) in predicates, such as would be required at some level for sentence (c) above, in particular (p. 184):

a. [Topic $\overline{X}$-wa] [Pred $\overline{X}$* V]

b. Topic Binding: The- topic $\overline{X}$ must be bound pragmatically to an X which is in the domain of Predication (Pred).

Kitagawa (1982:188) recognizes that (1) Topics may be PPs and adverbials as well as NPs, as in sentence (b) above; and (2) a sentence may have more than one Topic (with the possibility of a "contrastive" reading for the second Topic), as in:

Bunmeikoku-wa dansei-wa heikin zyumyoo-ga naga-i
civilized countries-TOP man-TOP . . . [as in (d)]

'As for civilized countries, as far as men are
concerned, the average life span is long[1]

Turning to other characteristics of J, PD is permissible in many environments according to primarily pragmatic criteria (cf. Hinds 1978). Serial verbs occur, although only in certain types of constructions; e.g., with ageru 'give,' morau 'receive,' kuru 'come,' and iku 'go. Cross-sentential Topic chains are common in discourse, and are related to the phenomenon of ellipsis in Japanese, which has received a considerable amount of attention (cf. Hinds 1978).

2.2.5 Mandarin

Mandarin is a Tp language in which the Topic function is "integrated into basic sentence structure, and Topic and Subject are distinct" (Li and Thompson 1976:485). It has Pragmatic Word Order (PWO), which means that the word order is quite flexible, and serves to convey pragmatic as

opposed to grammatical information (Thompson 1978). The basic sentence has an initial Topic, followed by a predicate whose CWO is (S) VO, with OV variation (Li and Thompson 1981). Mandarin is a language which permits Subject ellipsis (PRO-drop), and has no dummy Subject pronouns, and no true passives (only bei passives with an "accidental" connotation). Below we will elaborate on some of these features.

Topic in Mandarin is usually a definite NP which Occurs sentence-initially, and is separated from the Comment by an optional "pause or pause particle, and appropriate prosodic features. Preverbal position is reserved for definite referents, as sentence (a) below illustrates; indefinite reference is necessary when the nominal appears after the verb, as in sentence (b) (Li and Thompson 1978:228):

    a.   Zéi păo   le
            thief run ASP
            'The thief has run away'

    b.   Păo le   zéi
            run ASP thief
            'A thief has run away'

Comments often do not contain either a Subject or an anaphoric pronoun that is coreferential with the Topic NP. This is illustrated in the following sentence from Li and Thompson (1978:227):

    Nèikuài tián wŏmen zhòng dàozi
       that field we   grow   rice
    'That field (TOP), we grow rice (on it)'

Some TCs have the structure NP-NP-VP, in which the first NP is Topic, and the second is Subject; although they have different referents, there is an implicit whole-part relationship between the NPs. Examples from Barry (1975) follow; some of the constructions are obviously similar to J examples cited above (in Section 2.2.4).

    a.   Wŏ tóu téng
            I head   hurt
            'My head hurts'

    b.   Junggwó dì dà rén dwō
            China land large people many
            'China's territory is large and its people
             many'

c. Nèibán sywésheng (lǐ), tā dzwèi tsūngmíng
that class students (in) he most smart
'In that class of students, he is the brightest'

Although in the above three sentences there is a semantic whole-part relationship between the Topic and the Subject₁ this is not always the case in TC constructions; the following three sentences attest to this variablility in the relationship between Topic and constituents in the sentential predicate (Li and Thompson 1976:479):

a. Nèi-chang huǒ xīngkui xiaofang-duì lái de kuài
that-CLASS fire fortunate fire-brigade come quick
'That fire (Topic), fortunately the fire
brigade came quickly'

b. Dòngwu wǒ zǔshang bǎo-shoǔ zhèngce
animal I advocate conservation policy
'Animals (Top), I advocate a conservation policy'

c. Nèi-zuo fángzi xīngkui qù-nián méi xià xuě
that-CLASS house fortunate last-year not snow
'That house (Topic), fortunately it didn't snow
last year'

Sometimes the semantic relationship between the two nominals or constituents is locative or temporal (Barry 1975:6, Li and Thompson 1981:95, Xu and Langendoen 1985:5):

a. Jīntyan tyānchi hěn hǎu
today weather very good
'Today the weather is very good'

b. Chyángshang pája hěn dwō bihú
on the wall climbing very many salamanders
(Lit: The wall is climbing many salamanders)
'The wall has a lot of salamanders climbing on it[1]

c. 1968 nián 8 yuè 22 rì wǒ yǒngyuǎn bù huì wàngji
year month day I ever not will forget
'August 22, 1968, I will never forget'

d. Zài Táiběi kéyi chī de hěn hǎo
at Taipei can eat very good
'In Taipei one can eat very well'

Therefore, as Xu and Langendoen (1985) point out, both argument and non-argument positions can be topicalized in M; indeed, even verb phrases and entire clauses can be Topics according to Li and Thompson (1981:98-99):

37

a. Chū-qu   hē chá wǒ qǐng nǐ
    exit-go drink tea I invite you
    'Going out for tea, I'll invite (treat) you'

b. Lìshǐ-xì   kāi-huì   wǒ kěyǐ gēn
    history dept. hold-meeting I can with

    Lǐsì tí-yi-tí
    Lisi mention-one-mention

    '(When) the history department has its meeting₁
    I can mention (it) to Lisi'

Closely related to its classification as a Tp
language with PWO, M does not have dummy Subjects such as
English it or there. Thus, Li and Thompson (1976:468)
contrast the following M constructions with their English
counterparts:

a. Zhèr hěn rè
    here very hot
    'It is very hot in here'

b. Yǒu yī-tiáo māo zài huāyuán-li
    exist one-CLASS cat at garden-in
    'There is a cat in the garden'

c. Kěnéng zhè-chǎng zhànzhen jiù-yào jiēsu-le
    possible this-CLASS war   will-soon end-ASP
    'It is possible that this war will end soon'

The types of TCs that occur in M without Subjects are
instructive insofar as they reveal sources for the IL
constructions which are described in Schachter and
Rutherford (1979) and Rutherford (1983). Consider, for
example, the following M constructions (from Li and
Thompson 1981:88-89):

a. Nèi-běn shū chūbǎn le
    that-CL book publish ASP
    'That book, (someone) has published it'

b. Fángzi zào-hǎo le
    house build-finish ASP
    'The house, (someone) has finished building it'

c. Zhèi-ge tímù zuì hǎo búyào tí-chū-lái
    this-CL topic most good don't bring:up-exit-come
    'This topic, (you'd) better not bring it up'

38

Li and Thompson (1981:89) caution that

> it is important to be aware that Mandarin
> sentences such as [those above] are **not** passive
> constructions. They are simply **topic-comment**
> constructions in which the subject of the verb is
> not present.

Sentences with Subjects also occur in **Mandarin**, of
course, as in the **"double-Subject"** constructions
illustrated above. Subjects control **reflexivization**, but
the grammatical cohesion between Subject and Predicate is
very loose compared to English. **Li** and Thompson (1981:16)
explain that

> the subject is not marked by position" by
> **agreement**, or by any case marker$_1$ and in fact, in
> ordinary conversation, the subject may be missing
> altogether ▪ ▪ ▪

If Subject in M is difficult to define (other than by
its so-called "doing^n or 'being^n relationship with the
predicate), **"verb"** is possibly even more difficult to
define. For example$_1$ Tai (1982) and Li and Thompson (1981)
note that Mandarin verbs are not **coded** for tense and
agreement; **furthermore**, the **adjective/verb/coverb**
distinction is not a clear **one**, and copular linking verbs
are not used.

With regard to pronominal anaphora, there is optional
zero-pronominalization for Subject in simplex sentences and
in correlative structures (but obligatory PD in adverbial
clause constructions); for example **(Li** and Thompson
1978:263):

Yīnwèi Zhāng-sān xǐhuan nǐ, suǒyi (tā) lái zhèr
because Zhang-san like you therefore (he) come here
'Because Zhang-san liked **you,** he came here'

In terms of PD in M$_I$ **Li** and Thompson (1978:263) state that

> The typologically distinct characteristic of
> Mandarin pronominalization is that often an NP
> preceded by another coreferential NP is simply
> deleted. This deletion process may occur within a
> complex sentence or across sentence boundaries in
> discourse.

Another feature of **M,** referred to above, is that of
serial verb constructions [defined in Appendix **A)**, in which

39

a single Subject governs two concatenated verbs (Li and Thompson 1976:478):

Zhāng-sān mǎi piào jìnqu
Zhang-san buy ticket go-in
'Zhang-san bought a ticket {and went in/
                           to go in'}

In a second kind of serial verb (Li and Thompson 1976:478), the Subject is agent of one predicate (V) and experiencer of the other:

Wǒ huā le qián xiǎngshòu
I spend ASP money enjoy
'I spent money {and had a good time/
               to have a good time'}

A different kind of PRO-drop which is characteristic of Tp languages is discourse-embedded. In a series of sentences, we observe the degree to which a so-called "Topic chain" allows an earlier stated or established Topic to trigger the subsequent omission of coreferential Topics. Therefore, it appears that PD could be considered another effect of Topic Prominence (Gundel 1980, Gundel and Tarone 1983; but see Huang 1984). Pragmatic strategies are used by M listeners to retrieve the referential Topic. An example of a Topic chain from Li and Thompson (1978:264) follows:

a. Nèichǎng huǒ xìngkuī xiāo fáng-duì lái de kuài
   that fire fortunately fire-brigade come quick

b. Zhǐ shāo-le sān-ge fángzi, fàng-le yī-xie yān
   only burn-ASP three house, release-ASP some smoke

c. Wǔ-fēn-zhōng yǐhòu, jiù miè-le
   five-minutes later, then extinguish-ASP

'That fire (Topic), fortunately the fire-brigade came quickly. (It) only burned up three houses, and released some smoke. Five minutes later, (it) was extinguished.'

40

2.2.6 Summary of Part II

In this brief linguistic overview, we have tried first to define the constructs of Topic, Subject, Tp and Sp, although in doing so we have, admittedly, avoided addressing some of the thorny issues related to these. Secondly, we have presented the kinds of **Topic-Comment** constructions which are permissible in English, J, and M, and the extent to which they are considered either basic and unmarked, or derived and thus marked in the respective grammars. The types of constructions posited by Rutherford (1983) to occur in the IL data of J and M learners of ESL can be largely traced to **(1) L1** constructions and (2) the pragmatic strategies used by speakers of Topic Prominent languages.

In the next section, we consider the development of languages over time, from Topic Prominence to Subject Prominence.

2.3 Syntacticization of Topic

In Part I of this chapter, the phenomenon of transfer
was discussed in terms of the constraints that operate in
conjunction with transfer to make it possible for Ll
structures, discourse-functions, or typological parameter
settings to be either abandoned or retained. in IL.   It was
pointed   out that typological distance   or   learners'
perception of distance are not necessary or sufficient
conditions for transfer to occur.   Rather, the compati-
bility of the Ll or IL structure with other principles of
language development, markedness, pragmatic demands, and
universal   grammar serves to reinforce the   potential
transferability of a given structure or parameter.   For
this reason, it was seen that the production of Ll-like
topicalized constructions of various kinds is a natural
process in SLA.   However, the effects of the Tp constraint
are not only evidenced in the production of literal
translations of Ll TCs.   The same discourse function may be
reflected in the avoidance of some forms, such as relative
clauses, or the "overproduction" of some L2 forms which are
adopted to serve Ll discourse functions as well as their
usual function in the L2.   Thus, locatives and extra-
position in Japanese data, and existentials and heavy
Subjects in Mandarin data are considered to be
"overproduced" for this reason; this judgment is determined
by an IL performance analysis, rather than a comparison
with normative data.   Also, the relative pace of
acquisition toward Subject Prominence is yet another sign
of a transfer effect if, say, Japanese learners are able to
pass through the various developmental stages more quickly
than Mandarin learners, on the grounds that Japanese is
both Tp and Sp and has both pragmatic and grammatical word
order, and consequently is more sensitive to the Subject
function and position in ESL.

In Part II of this chapter, we reviewed Li and
Thompson's (1976) typological paradigm, and some of the
characteristics of Topics and Subjects, and then outlined
kinds of TCs which are permissible in the respective Lls.
English, it was noted, has Topic functions which are
integrated with those of Subject, and TCs of the Japanese
and Mandarin sort are considered relatively marked.   By
contrast, the range of TCs that can occur as unmarked
constructions in Japanese and Mandarin, such as double
Subjects in which the first element could not be derived
from any other element in the sentence, was highlighted.
Subtle differences between Japanese and Mandarin were not

42

expressly discussed, but it was pointed out that J has more Sp features than M and, furthermore, that it lacks some of the word order flexiblility and constructions such as serial verbs which are claimed for M.

In this section, we turn to a review of studies which have discussed the nature of syntacticization, as in the developmental shift from Tp to Sp, or in other words, from the "pragmatic" or "pre-syntactic" mode to the 'syntactic' mode. This overview will familiarize readers with the naturalness and origins of the constructions which Rutherford's developmental model posits for Japanese and Mandarin learners.

We will cite studies and main findings that address the fundamental question about the roles of Topics and Subjects in studies of language change and development, in universal grammar, and most particularly in SLA. We will proceed by outlining studies which look at various kinds Of developmental linguistic data, for as Givon (1979) and Zobl (1984) note, there should be consistency in what consti-tutes an "allowable" sequence in languages across all kinds of language development, here with regard to a shift from Tp to Sp. We will present the findings of child L1 acquisition studies, child SLA, adult SLA, pidginization, and diachronic studies of the syntacticization of Topics into sentence-internal constituents (primarily Subjects) at different points in time. This global perspective of the shift from Tp to Sp, coupled with an understanding of the various factors involved in the transfer of L1 parameters such as Tp, gives us a powerful and broad foundation upon which to make predictions and to analyze the J and M ESL data in the research portion of the thesis. For we shall see if the syntacticization of Topic in J and M ESL conforms to patterns established in other developmental data; and yet, within the established bounds of possible structures, how J and M might achieve the same processes differently.

2.3.1 *An* Integrated Perspective of Syntacticization

Givon's (1979) description, interpretation, and functional explanation of syntacticization from discourse to syntax (i.e., Tp to Sp), like Li and Thompson's linguistic typology presented in Part II, is central to the discussion in this thesis. For Givon, perhaps following Slobin's (1977) perspective, observes striking, non-coincidental parallels in (1) the diachronic process of syntacticization, or "the genesis of syntax ex-discourse"

43

(Given 1979:222); (2) the development of Creole languages from Pidgins; (3) child versus adult language; and (4) informal versus formal speech. A fifth area, that of SLA, was not specifically addressed by Givon (1979), but reflects many of the same developmental processes. Each of these facets of syntacticization will be considered separately below, but a general sketch of the similarities across the different types of language change will be useful here.

What Givon has observed in various kinds of data he has analyzed is that Subjects evolve from Topics by means of the development of grammatical and semantic coding conventions (see Section 1.3, above), such as grammatical agreement with verbal morphology, and the overall "tightening up" of sentence-internal syntax. Of major interest is that as TCs give rise to Subject-Predicate structure, word order no longer has a uniquely pragmatic function; rather, it takes on semantic and syntactic functions as well. Grammatical morphology increases; syntax becomes more complex; intonation and stress have a lesser functional role; the ratio of nouns to verbs within the sentence is greater as serial verbs develop into one main verb and various case markers, which are later bound to their respective nominal arguments; and verbs are semantically more complex (p. 223).

To summarize his position, Givon (1979:232) states that

> It seems rather clear that the pragmatic mode of human communication ▪ ▪ ▪ is ontogenetically and phylogenetically earlier, and, in terms of cross-language attestation, more universal than the syntactic mode. At the syntactic level, languages tend to diverge enormously. At the pragmatic level, they tend to be amazingly similar. The fact that this mode is always used under the stressful condition of no common language ▬ as in Pidgins or foreign-talk ▬ simply underscores its status as the universal common denominator.

2.3.2 Syntacticization in Diachronic Studies

Givon (1979) identifies several aspects of diachronic change related to syntacticization. First and foremost, Subjects evolve from Topics and grammatical agreement on the verb develops. This process is illustrated in Tok Pisin, a Papua New Guinea language (Givon 1979:210).

Another example of the reanalysis of sentence-external Topics to a sentence-internal equivalent (something akin to Topic/Subject/Focus) is found in Foley and Van Valin (1984). The authors trace the development of the clause-internal "Topic-Focus" constituent in Philippine languages, which in Tagalog is case-marked with the postposition ang. The position and function of the-ang-phrase is distinguished from a diachronically earlier sentence-initial and sentence-external Topic. Thus Foley and Van Valin differentiate Topics and "Pragmatic Pivots" (e.g., grammaticized Subjects, clause-internal Topic/Focus), the latter being a product of the syntacticization process in those languages. The English Subject is also analyzed as a Pragmatic Pivot in this work. A topicalizing movement rule still operates in Philippine languages, much as it does in English, by preposing constituents to sentence-initial position for strictly pragmatic purposes; this is no longer a basic construction in Philippine grammars though, as it seems to have been at one time.

Givon (1979:222-223) points out that all languages have available to them ways of making information salient for functional, pragmatic purposes; but in some languages, the pragmatic mode is more structurally basic, whereas in others, it is simply an optional alternative to a more syntacticized base structure (cf. Li and Thompson 1976).

I would like to posit two extreme poles of communicative mode: the pragmatic mode and the syntactic mode. In the cases surveyed above [i.e., of diachronic change], we have been dealing with the rise of the latter out of the former. That every language has both extremes -- as well as any intermediate in between -- [is] obvious.
(Givon 1979:222-223).

Other reflexes of diachronic syntacticization outlined in Givon (1979) include the evolution of: passivization (in e.g., a Bantu language and in Indonesian); relative clauses from conditionals (e.g., in Hittite; cf. Haiman 1978); subordination in the VP from coordination (e.g., in Arabic); causitivization and more complex verbs and NPs (e.g., genitives in Krio); cleft and WH-questions; sentential complements (e.g., Biblical Hebrew); and inflectional morphology.

45

2.3.3 syntacticization from Pidgins to Creoles

Another kind of diachronic change involves the syntacticization or creolization of Pidgins. Givon characterizes Pidgins as having: "no stable syntax" (p. 224); a slower rate of delivery; no systematic grammatical word order; "transparent" TCs, but little in the way of Subject-Predicate structure; and minimal subordination.

Creoles, however, evolve in the speech of children of Pidgin speakers, that is within just a few years, and they are more representative of a syntactic mode than a pragmatic one. The emergent syntacticized Pidgin languages (i.e., Creoles)

> possess the entire range of grammatical signals used in the syntax of natural languages, such as fixed word order, grammatical morphology, intonation, embedding, and various constraints.
> (Givon 1979:224)

2.3.4 Syntacticization in L1 Acquisition

Givon (1979) and Bates and MacWhinney (1979) illustrate how early child first language acquisition is characteristic of the pragmatic mode of communication, compared to later child language and adult language, which are representative of the full-fledged syntactic mode. Givon notes the similarity between early stages of acquisition and Pidgins, which he attributes to the speakers' focus on communication, lack of common pragmatic background, and the immediately obvious ("here and now") context of discourse.

Bates and MacWhinney (1979) and Zobl (1983) share the view that

> certain longitudinal research in several language communities supports the functionalist hypothesis that children use a combination of semantic and pragmatic factors to guide their discovery of surface grammatical devices. In addition, there is a certain amount of evidence to support the prediction that early pragmatic ordering will place comment before topic, while topic fronting will be discovered only after the child becomes aware of the fact that the listener's perspective is different from his own.
> (Bates and MacWhinney 1979:194).

46

Thus, there appears initially to be a stage with an ellipted Topic, and Comment only, followed by a Comment-Topic stage, next a Topic-Comment stage, and finally a mature syntacticized grammar. This evolution of Subjects out of Topics in child language acquisition was also described by Gruber (1967) in an early analysis of a two-year-old's production of CTs and TCs.

2.3.5 From Syntactic to Pragmatic Mode in Different Registers

Givon (1979) has emphasized the fact that the diachronic or developmental shift from a pragmatic language mode to a syntactic one does not entail the loss of the former. Rather, both modes are available to speakers of a syntacticized language, and when there is communicative stress or lack of formal constraints on language use, the adult speaker may revert to a more pragmatic mode. Here, we will consider first the production of TCs in informal as opposed to formal registers, and second, the role of topicalization in the "Foreigner Talk" register.

Informal registers are characterized by Givon (1979:229) as having more TCs and left-dislocations than formal (e.g., written) registers; a slower rate of delivery with more repetitions and pauses; simplification of grammatical morphology; and shorter verbal clauses.

Bland (1981) illustrates this point by examining informal, spontaneous English discourse, in which she reports speakers use a number of pragmatic strategies to introduce, retrieve, and echo discourse topics, as well as to repair communication breakdown. She claims that informal English allows TCs which are surprisingly very similar to those in Chinese, and examples she cites from her data include the following constructions (Bland 1981:33-34):

   a. Roy I didn't know very well then.
   b. Your pants, you can't go out like that.
   c. My papers, the whole break was ruined.
   d. My work, I'm going crazy.
   e. My family, well my mother, we didn't have
       even have a car for ten years.

Another case of the pragmatic mode being used is found in "Foreigner Talk," which is the speech addressed to those whose language proficiency appears to be less than native-like (Ferguson 1975). Again, a preponderance of TCs and

lexical, morphological, and predicate simplification are used to attempt to make the language more comprehensible to the nonnative addressee and more easily processed as well. This supports the notion suggested in the section on transfer, that TCS are psycholinguistically unmarked (although as Givon 1984 reports, TCs can be seen as marked in the sense that they disrupt the continuity- of discourse). Givon (1984:128) identifies the communicative effectiveness of TCs, in that

> If one has difficulty establishing a new topic, or if one suspects that the hearer is likely to experience such a difficulty, the most sensible strategy is first to make sure that the topic is firmly established, and only then to come up with the new information.

That this strategic simplification in fact aids communication and comprehension was the focus of Chaudron's (1983) psycholinguistic experiment. He examined the effect of ESL teachers' topic reinstatements on comprehension and recall by L2 learners, hypothesizing that in terms of psycholinguistic salience and pragmatic effectiveness, the ranking of topicalizing strategies would be as follows:

Rhetorical > Repeated Noun > If-clause > Simple > Synonym
Question   Noun

His hypotheses were supported to some extent, although it turned out that the proficiency level of learners determined to a large degree the topicalization effect.

In short, pragmatic strategies such as topicalization are used in languages and in certain registers of languages in order to reduce the linguistic and psycholinguistic complexity of discourse, and to thereby facilitate communication.

2.3.6 Syntacticization in Child SLA

In this section, a study of child SLA will be cited, followed by a review of adult SLA literature concerned with the development of syntax, especially of Subject, in ESL.

Zobl (1983) studied the ESL of young Francophone children partly in response to findings reported in the literature that children produce certain types of Topic-Comment (TC) constructions in early SLA (cf. Hatch 1976). One aspect of the study was to see how age might affect the

48

relative proportion of TCs found in the children's data, since previous research (e.g., Bates 1976) had shown that children between ages 4-9 years used a primarily syntactic N-V-N processing heuristic in L1 and L2 production, whereas younger children tended to use a pragmatic heuristic or strategy, as in the use of TCs.

In the ESL of his youngest French subjects, aged two to three years, Zobl found a large proportion Of both CTs and TCs, in addition to the basic SV(O) word order. This reflects the use of pragmatic strategies also evidenced in L1 production in children at that age. Examples he cites (1983:213) follow:

   a.   Ready the steak (CT)
         'The steak is ready to eat'

   b.   Soup give it to baby (TC)
         'This soup, I/we should give it to the baby"

   c.   Plane ∎ ∎ ∎ I go play (TC)
         'I'm gonna play with my plane'

Also found in the data were examples of zero Topic with comment (where the Topic was understood from context).

In the ESL data of older children though, namely four to five year olds, Zobl noticed a predictable degree of 'syntactic conservatism" (Bates 1976), whereby TCs were virtually eliminated from IL production. However, in the data of yet older children in other studies, Zobl found that there was a 'resurfacing of pragmatically based production,' such that Spanish ten year olds produced both VS strings and TCs. In addition to age, Zobl observed that the relative flexibility of word order in the L1, such as the pragmatic word order of Spanish, 'encourages a prolongation of the pragmatic strategies" (p. 218), since Spanish and French allow VS strings, and they appear in child L1 acquisition data for these languages, but not in first language acquisition of English. Recall the discussion of the transfer effect of PWO in the first part of this chapter (Section 2.1.5).

2.3.7 Syntacticization in Adult SLA

Next, studies which document the syntacticization of Topic in adult SLA will be presented. First, we will present examples from Schumann's (forthcoming) work, which attempted to determine whether TC (or "Theme-Rheme")

constructions were actually basic to the organization of pidginized varieties of English, Then we will look at Givon's (1984) description of three Pidgin-English speakers[1] low-level development toward the fully grammaticalized target.

2.3.7.1 Pidginized Varieties of ESL

The variability of Pidgins with regard to the basicness of TCs was seen in Schumann's conflicting results: in two subjects' data TCs were indeed basic, whereas in three subjects' data "listing" was favored over TCs. Initially, Schumann had found a full range of TC use in the nonsyntactic utterances of three subjects: from 6-92%. In his follow-up study, too, the same diversity of TC use prevailed. All subjects were naturalistic learners whose "basilangs" (early pidginization) evidenced pre-verbal negation, almost no verbal morphology, and from 23-58% non-syntactic utterances (i.e., utterances without verbs). One Chinese and one Mexican subject employed a majority of TC constructions in the corpus of non-syntactic utterances.

Examples from Schumann (forthcoming) follow. The Theme-Rheme classification is not defined clearly, but Schumann differentiates it from Topic-Comment in that the former, he claims, is a discourse-level occurrence, and the latter is sentence-level. The pragmatic functions are, however, the same in the sense that their role is to frame the following utterance.

a. You know Hong Gong ⁻
         <u>Theme</u>
                    [everything faydon, you understan']
                              <u>Rheme</u>

   [Eh eh ... buy everything ⁻ very easy]
                  <u>Topic</u>   <u>Comment</u>

   'You know, in Hong Kong, everyone has freedom, you understand. Eh ... to buy everything is very easy.'

b. And the school ⁻ three
   'And as for school, three (of the family) go.'

c. And me, ⁻ in nine xxx and nineteen fifteen ...
   'And as for where I live, it's building 1915 ...[a]

50

A Japanese subject and two other Mexicans used fewer TCs, though, and in their place employed what Schumann termed a listing or sketching strategy. Schumann recognized that the number of TCs produced was inversely proportional to the number of utterances a basilang speaker produced with verbs. In other words, the three subjects who did a considerable amount of listing in their narrative discourse also produced more verbs. Examples from these subjects follow. First is a narrative produced by a Japanese speaker describing her son and daughter-in-law.

a. [Speaking about her son]
b. Then nervous breakdown
c. Then garden work ...
d. Wife, got trouble, wife
e. Because, ah, he nervous broke-down
f. Time go in the hospital, you know
g. Then take money for bank, you know

Schumann interprets this as a list, but we might just as well reanalyze it and the rest of the narrative as a series of Topic chains (Hinds 1983, Chen 1984, Li and Thompson 1979) with PRO-drop, which are in part a product of the L1 Topic Prominence. Thus, after the speaker has introduced the Topic (her son) in the discourse preceding this text (a), she subsequently omits the coreferential Subjects (i.e., performs PRO-drop) in (b) and (c). Then, after the Topic shifts to the speaker's daughter-in-law in (d), there is at first a TC, followed by an utterance in (e), in which overt reference must be made to the son to avoid ambiguity. Thereafter, the remaining utterances (f) and (g) are understood as related to the already established Topic, the daughter-in-law. In short, Schumann might have captured the pragmatic strategy more effectively by observing its similarity with Topic chains used in, for example, Chinese and Japanese.

Topic chains and lists of predicates also occur in the two Mexican speakers' discourse, as do some TCs. Examples of TCs follow:

a. Me small ⁻ no scare
   'As for me when I was small, I wasn't afraid'

b. And then my daughter, Patricia ⁻ oh, right now, separate
   'As for Patricia, right now she's separated from her husband'

c. And me, ~ es, es, too hard my test.
'And for me, the test was too hard'

d. Water ~ No like
'As for water, I didn't want any'.

From the foregoing examples, it is apparent that in pidginized English, there is some ambiguity as to the source of TCs: whether L1-influenced, or whether fossilization has simply occurred in the pragmatic mode. Of course, we can claim that in the Japanese data, these two factors reinforce each other, a process which Andersen (1983) also found in the ESL no-Verb negation of Spanish speaking learners.

Givon (1984:129) posits that TC constructions are pragmatically basic in the sense that 'when communication is under severe stress, speakers of varying linguistic backgrounds revert to this common communicative mode." However, they can be considered more marked than constructions with the reverse order, CT, from the point of view of discourse continuity. That is, the most continuous and predictable discourse does not contain Topic shifts or overtly marked Topics, but rather, develops a Topic which is already known. By overtly expressing the Topic before the Comment, then, the utterance becomes marked.

In the transcripts he examined of Korean, Filipino, and Spanish learners[1] ESL data) Givon reports that the Korean data included such IL features as zero anaphora, a preponderance of definite NPs, left and right dislocated definite NPs, and Topic NPs without any predication.

The Filipino speaker) however, produced more pronouns than appeared in the Korean or Japanese texts Givon examined. He speculates, therefore, that 'some of the functional load carried by zero-anaphora in the Korean-English text is borne by (unstressed) pronouns in the Philippine-English text[n] (Givon 1984:121).

In the Spanish-English text, he found that there were numerous instances of Topic-marking by means of Topic repetition plus attempted comment.

On the basis of these observations and a cross-linguistic survey of Topic-marking in natural discourse, Givon (1984:126) concludes that a Topic-marking universal hierarchy (of least to most marked constructions) might take the following form, based on the predictability of

Topics and the shared knowledge that interlocutors are assumed to have:

           (1)   **(2)**
      zero anaphora > unstressed **PRO/** >
      [Comment only]   verb agreement
                          [Comment-Topic1

              (3)   **(4)**   **(5)**
      independent/stressed > full-NP > repeated full-NP
            PRO   **[T-C]**   [Topic only]
      [Comment-Topic]

To account for this hierarchy$_I$ he posits the following word order universals **(Givon 1984:126)**:

(i) <u>Quantity</u> **<u>Universal</u>**: More continuous, predictable, nondisruptive topics will be marked by less marking material; while less continuous, unpredictable/surprising, or disruptive topics will be marked by more marking material.

(ii) <u>Word</u> <u>Order</u> **<u>Universal</u>**: Of topics that are fully expressed as an independent word or pronoun, those that are most **continuous/predictable** will display COMMENT-TOPIC **(VS,VO)** word-order; while those that are less **continuous/predictable** will display **TOPIC-COMMENT (SV, OV)** word order.

Indeed, Givon (1984) hypothesizes that impetus for language change **(e.g.,** child **L1** acquisition and diachronic language change) is the relative markedness of **TCs** in discourse with respect to their unmarked counterparts, that is, those points which are **leftmost** on the hierarchy shown above. In short, **TCs** occur less frequently in continuous text than **CTs** or Comments only, and perhaps because of **this,** they have more pragmatic force.

Forerunners to **Rutherford's** Study

In **this** last part of the **chapter,** we present some of the studies of adult SLA in which a discourse-functional explanation for ESL IL data is derived from the Topic Prominent typology of the Ll. First, **Huebner's (e.g.,** 1979) work is described, and then an analysis conducted by Schachter and Rutherford **(1979)** in much the same spirit as Huebner's is outlined. It will become clear that these two studies were instrumental in the conception **of** Rutherford's **(1983)** analysis.

2.3.7.2 Huebner's Dynamic Paradigm

Huebner (1979, 1982, 1983, 1985), who is also interested in the Topic-marking strategies of non-native speakers of English, proposed a creative discourse-functional analysis of the ESL IL data of a Emong (Tp) speaker. In his work, Huebner observed the distribution of zero definite articles for Subject NPs, for which overt marking of definiteness was redundant for the speaker. This is because these sentence-initial NPs functioned as discourse (and sentence) Topics, and therefore? they had implicit definite reference. This Topic function (presumably transferred from the Ll, Hmong) was not coded with the English determiner the) rather, it was surface-coded with the IL postpositional Topic-boundary marker, is(a), which also functioned as the English copula (from which the form appeared to have been derived).

By studying his data from a "dynamic," functional perspective, Huebner was able to account for systematicity he found in the IL of his subject, Ge. Over the course of one year, Ge's production of overt Tp constructions, which had characterized his earliest IL, decreased and these gave rise to more syntacticized Topics; definite articles were supplied in the previously "zero" context, and the Topic-marking function of isa was dropped. A concurrent development in Ge's grammar was the acquisi-tion of referential pronouns in place of zero anaphora; dummy Subjects (non-referential pronouns) were not found in the data, however, and existentials were consistently constructed as (Ø) have NP.

Huebner's "dynamic paradigm" represents an insightful and rather revolutionary treatment of IL data (given the SLA research practices and procedures of the 1970s), especially insofar as it can explain the Hmong-English developmental phenomena in terms of Ll-influenced discourse-functional strategies. Other research which has examined IL data of adult learners from languages in which the discourse-syntactic notion of Topic figures prominently will be presented below.

2.3.7.3 Schachter and Rutherford (1979): A Look at J and M IL

Both the collaborated and independent efforts of Schachter and Rutherford have been directed at research questions and analyses that resemble Huebner's to some degree. Their specific focus has been different from

Huebner's, however, in that they have cross-sectionally examined the written production data of instructed ESL learners within the American university context.

For the purpose of teasing out plausible typological transfer effects in English, Schachter and Rutherford (1979) examined quantitative and qualitative differences in the ESL constructions of native speakers of Spanish, Arabic, Mandarin, Japanese, and Persian (Farsi) in written compositions. By comparison with an earlier study by Schachter and Hart (1979), which looked primarily at the production of complex clause-internal syntax by a similar assortment of subjects, Schachter and Rutherford (1979) adopted the typological parameter of Tp/Sp presented above to compare the data of their intermediate to advanced level learners. Their results showed a relatively high frequency of (1) extraposition in the J-ESL data, and (2) existential constructions in the data of M subjects. The overproduction of these sentence types was attributed to strategies J and M learners used to accomplish discourse functions with the L2 syntactic devices, it and there. In other words, the extraposition and existential usage was neither syntactically L1-like (since M and J do not have dummy Subjects) nor necessarily target-like and, moreover, was not shared to the same degree by learners whose L1s were Spanish, Arabic, or Persian as it was by J or M.

A closer analysis of the J and M data revealed that J used extraposition uniquely to introduce generic statements which would in turn function as future Topics; M, on the other hand, only used existentials to introduce new referents as Topics.

A few examples which appear in Schachter and Rutherford (1979) illustrate this main finding. In the first case, the future Topic is to be "friendly restaurants," and in the second, it is "tire."

(1) J- It is a tendency that such friendly restaurants become less in the big city.

(2) M- There is a tire hanging from the roof served as their playground.

A second area of L1 discourse-L2 syntactic crossover reflected in both J and M ESL structures is illustrated in the following sentences (from Schachter and Rutherford 1979):

55

(3) J- Most of food which is served in such
        restaurant have cooked already.

(4) M- Chiang's food must make in the kitchen of
        the restaurant but Marty's food could make
        in his house.

In previous accounts of the type of constructions found in (3) and (4), 'have cooked[n] and "must make[n] were often considered IL passives, corresponding to "has already been cooked[n] and "must be made,' respectively. In their reanalysis of the same, however, Schachter and Rutherford (who enlisted the insights of bilingual informants) interpret the sentences to be fundamentally TCs, in which the unspecified agent or Subject 'people[n] is deleted because it would be pragmatically recoverable in L1 discourse; and the object NP which is coreferential to the Topic (i.e., 'most of food," "Chiang's food," and "Marty's food[n]) is also deleted (see Section 2.2.5, above).

Following Schachter and Rutherford (1979), later works further elaborated on the strategies and L2 forms used by learners of ESL to perform L1- like discourse functions. Of particular interest were constructions which might be overproduced in IL for functional purposes, and, further- more, which might be accompanied by unique types of errors as a result. ∎

2.3.7.4 Rutherford's Pilot Study

Rutherford (1983) examined the interaction of several typological parameters in the written data of learners from the same L1 backgrounds as he had previously examined. That is, in addition to looking at Tp, he adopted Thompson's (1978) PWO/GWO typological paradigm and the more commonly known canonical word order paradigm (CWO). His working hypotheses, based on the Schachter and Rutherford (1979) study, were that in M-ESL, there would be frequent serial verb and existential constructions as the sentence below illustrates (Rutherford 1983:360):

    M- There's alot of people find their husband or
        wife in parties [note 3]

Also, as in sentence (4) above, more "surface" passives (which are really to be viewed as TCs with unexpressed agents) would be produced by M and J learners (and the former in particular) than by learners of non-Tp languages,

such as Arabic. Indeed, these two hypotheses were supported.

An unpredicted finding, however, was that M produced more sentential or 'heavy[n] Subjects than other learners in introductory topic sentences. Rutherford interpreted this last result to be in part a function of the length of the titles of the assigned writing tasks: i.e., heavy Subjects occurred more often in compositions whose titles were sentential, and especially for lower level learners, than for titles which were simple NPs. Examples of these opening sentences follow, and they were generated by the sentential composition title "How a man or woman chooses a wife or husband in my country[n] (Rutherford 1983): the levels of learners who produced these is also indicated, in ascending order of proficiency, although no further description of proficiency levels is provided:

a. A man chooses a wife is a man's business (level 3)

b. Of course his honest or not, self-confident or not, genteer manners or not, is very important (level 4)

c. Choosing a husband or wife in my country is quite a interesting thing. (level 5)

A more functional explanation, or hypothesis, posited by Rutherford (1983:361) was that:

> this frequent occurrence of so-called "heavy" subjects in the written English production of Mandarin speakers was evidence of a direct influence from the mother tongue, whose topic-prominent typology is strong enough to override more general acquisitional strategies that limit the early production of such constructions. [note 4]

On a production continuum, Rutherford noticed that at one end, while lower proficiency (Mandarin) students generated a number of sentential Subjects, they were qualitatively different from constructions at upper proficiency levels (as in sentence (c), above). The advanced students were able to use alternate, more syntacticized (and thus more native-like) constructions such as infinitives and gerundive nominals, and relative clauses with resumptive Subject pronouns.

2.3.7.5 Rutherford's Model

Next, Rutherford (1983) tested his hypothesis that Tp speakers, and particularly M, tend to produce heavy Subjects for reasons given above. Furthermore, he hypothesized that M would also characteristically produce serial verbs and TCs. In order to analyze the predicted developmental shift in sentential Subjects, he had learners write topic sentences for a set of six assigned composition titles; it was thus a simulated composition task. His subjects included 29 M, 39 Korean and J (combined), 66 Arabs, and 20 Spanish, all at various levels of ESL proficiency.

The results revealed that M produced significantly more TCs than other language groups. Secondly, while not statistically significant, due perhaps to the small number of tokens, the trend for M to produce more serial verbs than J was clear, reflecting the L1 influence. Thirdly, the trend for M to produce more sentential Subjects than the other language groups was also confirmed, but again, not statistically.

To explain this developmental paradigm for M, Rutherford (1983:363-364) devised a model representing the progressive syntacticization of a Tp grammar to an Sp one. This model has six phases corresponding to approximate and only hypothetically discrete stages of increasing proficiency. M learners gradually map onto Topic the syntactic functions of Subject, and onto Comment the non-sentential predicate structure of English, a process which Rutherford conceptualized as in Table 2.1 below; the accompanying examples come from the data in his study.

Table 2.1

Rutherford's Model of Mandarin Development

(i)
```
     T    -    C
   [NP ]      [Ø] VP
      T         C
```

e.g.,   Take good physical care of themselves is very important.

(ii)
```
     T    -    C
   [NP ]      NP  VP
      T          S-P
```

e.g.,   Alot of people, they know how to take good physical care by themselves.

(iii)
```
     T    -        C
   exist          VP
   [NP     ]
      T/S          P
```

e.g.,   There are a small amount of people get married in their teenage.

(iv)
```
     T    -    C
   exist      to V
   [NP    ]
      T/S       P
```

e.g.,   There are many elements to maintain a successful marriage.

(v)
```
     T    -    C
   exist      Rel Cl
   [NP   ]
      T/S-P
```

e.g.,   There are many problems that can make marriage unsuccessful.

(vi)
```
     T    -    C
   [NP   ]    VP
      T/S      P
```

e.g.,   More people do physical exercises now than before.

59

In this dynamic paradigm shift from Tp to Sp, we also
see the possibility for IL PRO-drop at stage **(i)**, and
later, the introduction and evolution of existential <u>there</u>
constructions, particularly at stages **(iii)-(v)**.
Rutherford notes that between stages **(iii)** and **(iv)** where
one might expect to see a theoretically possible
construction such as **"There** are a small amount of people
<u>they</u> get **married,"** this evidently does not occur. We
observe the avoidance of early relative clause production
as Schachter **(1974)** predicted, and a kind of restrictive
simplification **(Meisel 1980)** at stage **(i)** where the second
NP is deleted, which later gives way to an elaborative
simplification with the redundancy shown in **(ii)**, with the
pronoun <u>they</u>. Again, there is an abundance of existential
<u>there</u> constructions fulfilling the function described
above, and which **Hilles (1985)** also identified as a
possible structural pivot in the syntacticization process.

With regard to J production data, Rutherford roughly
charted the occurrence of sentence-initial locatives which
he interprets to be the developmental source for existen-
tial <u>there</u> constructions or extraposed sentential Subjects
with <u>it</u>. The J paradigm is presented by Rutherford **(p.**
365) as follows in Table 2.2.

Table 2.2

Rutherford's Model of Japanese Development

(i)   T -   C
          loc   VP
          in NP

e.g.,   In my country, hasn't army, navy and air
        force.

(ii)   T -   C
          loc   VP
          in NP   there - BE
                  or NP

e.g.,   In America, there are many kinds of people.

(iii)   C   -   T
        There BE VP - loc
                    in NP

e.g., Because_r there are many Japanese in U.S.C.
....

By way of final summary and comparison of the Mandarin
and Japanese developmental ESL constructions, Rutherford
(1983:367) observed that

> The commonality side was demonstrated by an aspect
> of the syntacticization process: interlanguage
> progression from topic-comment to subject-
> predicate in the acquisition of **sentential**
> subjects by Mandarin speakers and in the
> acquisition of existentials by Mandarin, Japanese,
> and Korean speakers. The differential side was
> revealed in the extra-heavy topic-comment
> influence from Mandarin and the grammatical **word-**
> order sensitivity from Japanese and Korean.

Thus the quantitative differences in the J and M data
are that J had a tendency to produce more dummy Subjects
than M. (Note that this contradicts Schachter and
Rutherford's 1979 finding that M produced more <u>there</u>
pronouns than **J,** but J more <u>it</u> pronouns than **M.)** The
difference in J and M production in Rutherford **(1983) of**
<u>there</u> was significant, but it was not significant for <u>it</u>.
The qualitative difference was that there were no serial

61

verbs in **J** data, and that locatives seemed to play an important developmental role for **J** ESL learners.

A finding in both corpora of data was that the typological parameters of **Tp/Sp** and **PWO/GWO** were important in determining plausible L1 discourse-functional transfer effects; CWO, however, was judged by **Rutherford** to be relatively unimportant. Although Rutherford describes these developmental stages and cross-language differences with some empirical grounds (methodological snags aside; these will be discussed in Chapter III), he does not speculate on the issue of what kind of linguistic input in the **L2** might be necessary to bring about the learners' progression or about how the development of Subject-verb agreement fits into his model. We might speculate, however, that as we see a gradual syntacticization of the **VP** node and the types of operations governed by Subject, the following sorts of information in the English input provide the "positive evidence" that is required: **(1)** NP-VP agreement; **i.e.,** surface-marking of third person "-s," which signals the basicness of the grammatical cohesion between Subject and **VP;** (2) the appearance **of** dummy Subjects; **(3)** the fact that English is not a PRO-drop language and, although a coreferential PRO (Subject) in an infinitival or coordinated clause may be dropped and there is some ellipsis of Subjects in informal discourse, a relative pronoun is required by. stage **(v).**

For **J** learners, **it** appears that in the last of three proposed stages, the locative **(PP)** occurs in sentence-final position. This suggests that the commonly **topicalized,** sentence-initial locative adverbial in Japanese, which serves as a frame or scene-setter for the rest of the sentence (Smith 1978, **Kuno 1973),** is gradually **syntacticized** as a VP-internal PP in ESL. The acquisition of adverbial position is considered to be problematic for **J-ESL** learners even at advanced levels **(cf.** Zobl **1985a).** This is perhaps due to a combination of several factors: **(1)** the **L1** influence referred to above; **(2)** the pragmatic preposing of adverbials that is also possible in English: (3) difficulties with learning sentence-internal adverbial position; and (4) problems posed by English preposition usage. Thus, even though the adposition parameter **(i.e.,** pre- versus post-positional) is acquired early, on account of the apparent complexity **of** lexical choice **of** preposi-tions, learners may use a strategy to avoid prepositional phrases by preposing the NP and marking **it** with either a zero preposition or the set (basically formulaic) preposition <u>in</u>. This strategy is confirmed by the

introspective data of a Japanese informant that was presented at the beginning of Chapter I.

2.3.7.6 Curran (1984)

Turning now to study by Curran (1984) of Korean-ESL, we find data that on the one hand are consistent with Rutherford's observations of the Japanese learners' IL (since both Japanese and Korean are attested TSp languages), and which, furthermore, support a position presented earlier regarding the negligible role of transferred canonical word order in all but primitive SLA. Curran found that the production of oral narratives by her two native Korean subjects evidenced a large number of putative TC constructions and left-dislocations. There was also a rather large proportion of sentence-initial temporal and locative adverbials (sometimes followed by the Topic marker is, as in Huebner 1979), which also served the discourse-functional purpose of Topic. Curran estimated that one quarter of the sentences in the corpus of data she examined contained TC features, again testifying to the transfer not of L1 CWO (to which Curran's primary research question was addressed), but of pragmatic strategies for framing utterances and identifying known referents.

For example, her first subject tended to establish temporal reference sentence-initially with an adverbial, as in the following sentence (Curran 1984):

(1)   In that time is I call the towing car.

This preposing of adverbials was identified as an L1 transfer effect, based on Hwang's (1985) observation that Korean, like Japanese, places topicalized adverbials with optional locative case markers in sentence-initial position.

Curran's second subject, by contrast, used a number of generic Topics, partly due to the philosophical nature of the narrative he was giving. For example:

(2)   That's   why people they   eh   lifetime   only mental practice.

Curran felt that the two subjects were somewhere in between Tp and Sp in the syntacticization process, and that the factors which seemed to be *reinforcing[n] each other in the IL production of TCs were: (1) the functional role of Topics   in   terms of   contrast;   (2) simplification (as   in

copula deletion in **existentials**); and (3) topic **clarifi-**
cation, usually involving quantification (**p.** 1O-11)**.** An
example of the overlapping TC functions follows:

> Interviewer: How many children .are in your
>   classes?
>
> Subject: Korean class ... twenty-six, twenty-six
>   kids,   And then American class uh
>   twenty-four kids.

2.3.7.7   Rutherford and Altman (1985)

Rutherford and Altman (**1985**) undertook a further
analysis of **IL** data with particular interest in the notion
of "discourse competence" in a second language. The
authors attempted to determine whether

> in the early stages of learning to use the target
> language, learners will fall back on those aspects
> of communicative competence acquired in their
> native language that can be put to use in the
> second language as well -- **i.e.,** that can be
> "transferred".
>
> (Rutherford and Altman **1985:2**)

By communicative competence, they **refer** to the use of
Ll PWO using **TCs.** They hypothesized that CWO, namely
syntactic word order, would not transfer. They stated that
their hypothesis would be confirmed if, say, speakers of a
PWO and SVO L1 (**e.g.,** Spanish) produced SOV in ESL **IL;** or,
if subjects of an L1 with GWO but not SVO (**e.g.,** Japanese)
produced SVO in IL, reflecting the sensitivity of speakers
of a GWO **Ll** to GWO in the **L2.**

Rutherford and Altman examined a total of 300 writing
samples such as were described in other studies (Schachter
and Hart 1979, Schachter and Rutherford 1979, Rutherford
1983). The Ll groups involved were Japanese, Arabic, and
Spanish. The results were that J produced all SVO strings
in **IL,** as predicted. The Arabic and Spanish subjects, on
the other hand, produced XVS strings, but there was no CWO
transfer (**e.g.,** VSO for Arabic). Therefore, the claim that
PWO and Tp are more likely to'transfer than CWO, at least
at the level targeted here, was borne out.

64

2.3.7.8 Duff (1984)

Following the discourse-functional analyses of Huebner (1979), Schachter and Rutherford (1979), Rutherford (1983), and Curran (1984), and with particular interest in testing the transfer of the Tp/Sp Parameter in SLA, Duff (1984) collected conversational production data from a total of sixteen J and M learners of ESL (i.e., eight learners from each Ll). Based on the previous research, Duff hypothesized that: (1)M would produce more TCs in ESL than J (theoretically due to the more basic nature of Topic in M than in J); and (2) there would be qualitative differences in the kinds of constituents topicalized across the two languages; e.g., sentential Subjects for M, and adverbials (e.g., locatives, temporals) for J.

The subjects in the study wete students at the English Language Institute at the University of Hawaii, whose TOEFL scores were, on average, from 450 to 500. For the most part, subjects had lived in Hawaii for less than three years. The topics discussed by the subjects were based on controlled problem-solving tasks and debates, which were performed together by subjects in dyads.

There were three types of constituents regularly topicalized in the data: (1)NPs, (2) sentences (whole or elliptical), and (3) adverbials (e.g., locatives and temporals). In accordance with Duff's hypotheses, she reported the tendency for (1) M to produce more TCs than J, and (2) for M and J to evidence approximately the same proportion of topicalized NPs; M, however, produced a greater number of sentential Topics than J, and relatively few topicalized adverbials. The J data revealed exactly the opposite pattern, with adverbials accounting for a higher percentage of the Topics than heavy Subjects. There are, admittedly, some methodological problems with the nature of the quantification of the relevant data, due to the lack of a baseline unit of measure by which to compare J and M output (especially as concerns the first hypothesis). Nonetheless, the results (the qualitative differences in particular) suggest some interesting patterns.

Examples of NP, sentential, and adverbial Topics in J and M data, respectively, are presented below.

Japanese

(1) Their thinking, because who lives uh more for fifty or sixty years, so seeing is believing.

(2) But I think matches, is for ˙ like cooking something.

(3) First aid kit, what's for?

(4) But what if I have television, uh we can get uh good English and convenient.

(5) So he didn't drink much, so he didn't have za accident.

(6) If we have the children, so we cannot watch the news show and Chinese language program on TV.

(7) But present society-in present society, So everybody criticize society and political things.

(8) Desert island, if you carry the fresh fruits, it's soon dried.

(9) You know in Japan, we have za educational program.

Mandarin

(10) Yeah the people, you mean it's less talk with their family?

(11) And sleeping bag, go to sleep.

(12) And then children, he just learn from TV and do something.

(13) What do you mean, I don't know.

(14) But this compare with the older and the younger, I think they're the same.

(15) Not because you experience you will be bad, I don't think so.

(16) In China, in the city got this kind of trouble.

(17) Yeah so <u>in the daytime, from morning from eight to eleven, and</u> afternoon <u>from two to five,</u> is the TV university.

(18) <u>Only in the morning and in the nighttime after six news, and then until eleven,</u> isa entertainment.

In the whole corpus of data, as illustrated by the above examples, it is noteworthy that dummy Subjects such as <u>there</u> and <u>it</u> scarcely occurred. Compared with, say, <u>written</u> production data reported in Schachter and Rutherford (1979), Rutherford (1983), or Rutherford and Altman (1985) from a roughly comparable group of learners. Instead, in the M data, for example, alternative ways were used to express information which might otherwise occur with expletives; in (19) below, the empty preverbal Subject position is filled with the patient role <u>he</u>, not <u>it</u>, whereas the impersonal Subjects in (20) and (21) are expressed with the generic pronouns <u>you</u> and <u>we</u>, plus the verbs <u>have</u> or <u>get</u>. This might be indicative of the fact that these utterances correspond to about stage (ii) in Rutherford's (1983) model in Table 2.1 above.

(19) According to this case, yeah, because the car driver₁ he can help him down an then he take x hospital sooner, maybe <u>he don't necessary to cut off his leg.</u>
('. . . if the driver had stopped to take the injured man to the hospital, it wouldn't have been necessary to cut off his leg')

(20) In Chinese culture, <u>you</u> have the <u>expect to the oldest</u>.
('. . . it is expected that . . .')

(21) If <u>we</u> <u>not</u> <u>have</u> <u>TV</u>, we also can get international communication.
('Even if there were no/wasn't any T.V., we could have/there would be international communication')

One might, in this same light, consider the sentences in which the adverbial-fronted existentials with inversion (and minus PRO), e.g., (16) to (18), are representative of an earlier stage than sentences (19) to (21) (Rutherford and Altman's 1985 data shows the same tendency). These also support Zobl's proposed "one constituent constraint," according to which Subject PD is greater where lexical

material (a syntactic constituent) other than the Subject occurs in preverbal position, and especially where a non-referential pronoun would be expected (Zobl 1984:197). Examples Zobl (1984:202) cites from the Spanish speaker Alberto's data are:

a. In England is more big the problem.
b. For me is better the beer
c. In this country is crazy the weather, no?

Returning to the Duff (1984) data, in addition to the relatively low proportion of extraposed or existential constructions of the expected type, there were several other noteworthy features which required further investigation of the functional sort of analysis that is currently advocated in the literature. First of all, an unexpected finding was that <u>conditional</u> constructions appeared far more often than predicted, and this signalled the possibility of a discourse–syntactic explanation (supported, moreover, by Haiman 1978). Rutherford (1983) made only passing reference to the appearance of conditionals in his data, namely that there were more <u>if</u> and <u>when</u> constructions in J-ESL than in M-ESL, but he did not elaborate. ∎

A classic$_1$ rather complicated example from the M data in the study follows in (22), which illustrates the kind of conditional usage that is being referred to:

(22) But I say you should compare with <u>if I for example</u> I <u>live in America</u>, so everyday I just look something and some idea expressed from America ∎∎ But <u>if I</u> didn't <u>receive, if I</u> haven't <u>seen any program compared</u> these <u>two</u> situation, I think you see something is better than you see nothing.

In (22), there are two conditional expressions, the first marked with <u>if . . . so,</u> and the second consisting of the elements <u>if . . . if . . . [zero].</u> A number of other comparable constructions occur in the sentences above, (namely 4, 5, 6, 8, 11, 14, 19, 21) and they are typically coded with various combinations of the markers <u>if, so,</u> and zero (0).

One plausible interpretation of this large number of conditionals (i.e., "overproduction") is that they are fulfilling a TC function$_1$ with the <u>if</u> clause as Topic or topicalizer of information which in the tasks used is more or less *given^n (sometimes counterfactually, in which case the fact is given in the discourse; see discussion in Section 2.2.2). Conditionals are, thus, a subset of the

sentential–Topic category, which behave in much the same way as other Topics do; i.e., they provide a framework for the comment, and presupposition on the basis of which to make new assertions. [note 5] Similar examples of if-clause (but non–conditional) topicalization produced by ESL teachers in Chaudron's (1983) study were as follows (pp. 442-443):

    a.   If you look at the amount of rain (Topic),
        if it's under 20 inches, that's not very much rain.

    b.   If you look at the Sun [newspaper] (Topic),
        or something, you'll notice that every sentence is a paragraph just about.

Besides the role of conditionals in Duff's J and M IL data, a second point which deserves further attention is the role of agreement between or among Topic, Subject, and verb. If indeed Subject–verb agreement is posited as a trigger in the syntacticization of TC sentences, or at least a necessary milestone in the same process, it should be incorporated in the analysis as evidence for development in this direction. In sentences (8), (10), and (12), a generic, plural NP is Topic, e.g., people, followed by a singular coreferential pronoun (e.g., unanalyzed it's). Presumably, this stage would be preceded by sentences with an ellipted coreferential PRO (e.g., as in 2, 3, 11, 16), and followed by the analyzed third person conjugation (as in 14), and at more advanced levels, the generic Topic might be replaced by the specific (perhaps singular), syntacticized Subject and would be accompanied by Subject–verb agreement.

Interestingly, as mentioned above, the coreferential Subject pronoun it/it's is not deleted (i.e., by PRO-drop), although the PD distinction between J and M (+ PD) and English (-PD) is, arguably, a typological parameter which could very easily be transferred in these environments (White 1984, 1985, Hilles 1985). Thus the filler PRO it's hints of an intermediate stage between the "pre-syntacticized" and "syntacticized" modes, and would thus be represented as stage (ii) in Rutherford's model for M learners (see Table 2.1, Section 2.3.7.5). whether the same usage would occur in written production is not clear, although the following chapters will report a study analyzing the written data of a comparable group of learners, and this question will be addressed there.

69

2.3.7.9 Summary of SLA Studies

In the numerous studies cited **above,** the repeated occurrence of **TCs** was observed in the **ILs** of learners whose **Lls** were Tp and **TSp,** such as Hmong, M, **J,** and Korean$_I$ and whose IL CWO was (in all but the Pidgin data) invariably SVO. The literature described the intersection of **Ll** Tp typology with L2 Sp typology, in terms of <u>direct</u> and <u>indirect</u> effects, as in the **"overproduction"** of **sentence-** initial adverbials, nominal Topics, and **if-clauses,** in the data of learners of approximately intermediate proficiency levels **(e.g.,** TOEFL **450-500).** The following points **summar-** ize the discourse-functional effects in the **ILs** of the subjects in the respective studies:

(1) **An** IL article system based on the implicit definite reference of Topics; PD at early developmental stages **(Huebner 1979);**

(2) A relatively high frequency of sentence-initial adverbials in Japanese- and **Korean-ESL,** reflecting a kind of topicalization by means of locatives and temporals commonly used to this end in the Ll (Rutherford 1983, Duff 1984, **Curran** 1984);

(3) A large number **of TCs** with Topic **NPs,** sometimes coded in IL with the postposed marker <u>is(a)</u> **(Schumann,** forthcoming, Givon **1984,** Huebner 1979, **Duff** 1984, **Curran** 1984);

(4) A subset of **TCs** in (3) with sentential Topics, especially in M data (Rutherford **1983,** Duff **1984);**

(5) The functional role of extraposition in J and existentials in M to introduce new Topics (Schachter and Rutherford **1979,** Rutherford 1983):

(6) The role of conditionals **(e.g.,** <u>if</u> clauses) as **Topics/topicalizers** (Duff 1984).

In the early SLA (Pidgin) data of Japanese, Mexican$_I$ Spanish and Filipino subjects **(Schumann,** forthcoming$_I$ and Givon **1984),** the same pragmatic strategies leading to the production of **TCs,** Topic chains and listing were also observed, although possibly for different reasons. The Pidgin data characterizes the **pragmatic** mode, which appears to be the result of fossilization at a level reflecting the stressful multilingual context in which migrant workers typically find themselves. Hence, the potential transfer effect **of** Tp and/or **PWO** is reinforced for these learners,

70

due to communication demands and the frequency of similar structures (e.g., TCs) in the input to them.

2.3.8   Summary of Chapter II

In the foregoing sections, we considered ways in which syntax develops out of discourse-functional strategies and in many kinds of language change. Specifically, we observed that Subjects evolve out of sentence Topics as VP agreement is acquired and Subjects adopt appropriate linguistic behavioral and coding properties.

Theoretical and experimental analyses of this process were reviewed and explanations were recounted for the kinds of developmental stages that Rutherford (1983) posited would lead up to full syntacticization. Hence, our survey not only examined SLA data, but also manifestations of the same phenomena in diachronic, ontogenetic, and creolization data.

The universally attested pragmatic mode in early language development gives rise to the syntactic mode, and this change is reinforced by principles of markedness theory, perceptual processing constraints, and core grammar, among others. Although syntacticization is seen to be a natural developmental process, attestation of the same process in J and M learners' ESL maturation also reflects the prolonging influence of pragmatic strategies, which are partly grammaticalized in the learners' L1s. Furthermore, in a form-to-function analysis of J and M IL data, discourse-functional effects revealed that Subjects evolve from existentials in M-ESL, and from locatives and then existentials in J-ESL. The development of relative clause formation is yet another aspect of syntacticization.

In the following chapter, research conducted by this author in response to findings and hypotheses reviewed above will be outlined.

71

CHAPTER **III**

THE STUDY

3.1 Purpose

The purpose of this study is to verify and perhaps refine the claims made in the literature by Rutherford **(1983)**, in particular, regarding the transfer of **Ll** discourse-syntactic processes in the ESL production of J and M learners. Although the present research is different from previous work in that it deals exclusively with Japanese and Mandarin **ILs** from the outset, we have tried, as much as possible, to ensure cross-study comparability by using the same kind of target population, task type, and modality as in Schachter and Hart **(1979)**, Schachter and Rutherford **(1979)**, Rutherford **(1983)**, and Rutherford and Altman **(1985)**. Furthermore, the present study improves upon the previous work by being explicit about research methodology, by performing random sampling, using baseline units of measurement, reporting inter-rater reliability, performing inferential statistical analysis of the strength of relationships between the variables, and so forth.

3.2 Research Questions

The primary research question is: What is the effect of the independent variables of Ll (**J** or **M**) and proficiency level on the dependent variable of syntacticization of Topic in ESL? Secondly, to what degree are these effects accurately captured by Rutherford's model **(1983)?** Thirdly, how is the claimed transfer effect of Topic Prominence consistent with other constraints on language development, such as psychological operating principles, markedness theory, psychotypological factors, and UG?

3.3 Hypotheses

The hypotheses are based on the research findings reported in the latter part of Chapter **II.** The null hypothesis, which predicts no differences between **Ll** groups on the various dependent measures, would be compatible with the view of second language acquisition that attributes all language development to learner-internal and **L2-** internal factors. Because the research outlined above claims just the opposite, however, namely that there are

L1-induced processes or constraints on L2 development and production (in conjunction with natural developmental sequences and principles of universal grammar and markedness theory), which result in differences in production, we will choose alternative hypotheses. While a conservative approach would simply predict differences in J and M IL constructions, we will venture to assume a stronger position, one that predicts directional differences in the data. The hypotheses are listed below.

H (1): M learners produce more TC constructions than J (i.e., at stages (i) and (ii) in Table 2.1) due to the greater Tp of M vis-à-vis J, and the relatively higher frequency and flexibility with which TCs occur in M.

H (2): Among the TCs referred to in (1), M also produce more sentential or "heavy Subjects[n]" than J.

H (3): M produce more existential constructions than J, marked with the non-referential element there. This hypothesis differs slightly from Rutherford's claim that J produce on the whole more dummy Subjects than M (which, in any case will be tested in Hypothesis 5). However, according to Rutherford's model, we should expect more existen-tial development (as in his six-stage analysis) from M, considering, firstly, the results of Schachter and Rutherford (1979) which pointed to this, and secondly, the claimed functional role of existentials in M-ESL.

H (4): J produce more extraposition constructions (marked by non-referential it) than M.

H (5): The combined total of dummy Subjects (it and there) is greater for J than M. This prediction is based on Rutherford's (1983) notion that J, as a partial Sp and GWO language, sensitizes learners to the production of non-referential place-holders in ESL.

H (6): J produce more sentence-initial locative constructions, with or without the existential there, than M [at higher proficiencies, the locatives occur postverbally following an existential construction, but this will be tested in (10), below]. This is implicit in

Rutherford's model, as shown in Table 2.2 (Section 2.3.7.5).

H (7): J produce more "trueⁿ passives than M, due to the partial Sp and GWO status of their L1, and the wider range of passives in J than in M.

H (8): J produce more sentence-initial **if** constructions than M, and this reflects a discourse-functional strategy of encoding presuppositions and framing assertions (see Duff 1984, Haiman 1978).

H (9): J produce more **when** constructions at the beginning of sentences than M, for the same reason as given in (8).

H (10): There are differences in the effects of Tp, e.g., the amount of TCs, based on L2 proficiency. Thus, as learners syntacticize their constructions, they consequently produce fewer constructions at the lowest levels of Rutherford's model, fewer other constructions with the same discourse function (e.g., conditionals), and more Subject-verb agreement.

Other hypotheses which follow directly from the previous research but which are not central to the present study are that M produce more PRO-drop, serial verbs, and Topic chains than J.

3.4 Methodology

3.4.1 A Critique of Methods in Previous Studies

In several places in the preceding pages, it was suggested that the studies reviewed were weakened by vague descriptions of research methods, and sometimes serious problems threatening the validity and reliability of the research. In this section, before discussing the methodology used in the present study, a critique of those previously used is warranted. Since we are concerned primarily with cross-sectional studies with written production data, we will reserve our comments to Schachter and Rutherford (1979) and Rutherford (1983), (although Rutherford and Altman 1985 suffers from some of the same imprecision).

In reviewing the methodology of the aforementioned studies, the following points will be addressed: **(1)** sampling; **(2)** n-size; **(3)** proficiency levels; **(4)** baseline units of measurement; **(5)** validity; **(6)** inter-rater reliability; and (7) generalizability on the basis of inferential statistical analysis.

First, the selection of subjects (**i.e.,** compositions) in both studies is not explained, and it is not at all clear that there was random sampling. Second, the sample size is not clearly defined; for example, Schachter and Rutherford had at their disposal **525** compositions, but it is not stated **(1)** how many J and M were included in the total sample; (2) how and why they chose to analyze all 525 compositions for **extraposition,** but only 100 for existentials; or (3) how they could compare J or **M** with 'other groups," in terms of averages of the selected measures, as in extraposition in 3 out of 4 compositions for **J,** but 2 out of 4 for "others[n]. Does "others[n] here represent an average of all other groups pooled together, with the risk of squashing meaningful differences? In Rutherford **(1983),** similarly, the sample size is never clearly stated or sampling procedures explained. **We** can, however, calculate that on the basis of topic sentences produced on a six-sentence task, the groups include 29 **M,** 39 **J** and Korean (together), **20** Spanish, and **66** Arabs. The n-size, therefore, is rather small, especially since only six sentences are examined for each subject; it is difficult to make the kinds of generalizations Rutherford has posited, on this basis.

Third, the proficiency levels of groups are not explicitly provided, either in terms of local (University of Southern California) test scores or class **placement,** or in terms of standardized test scores, such as TOEFL. Additionally, there is not even a hint of the number **of** students at any approximate level. In other words, in a model such as **Rutherford's,** in particular, the development of syntax must be correlated with some kind of index of proficiency, but this is done only in very general terms, and hypothetically at that.

Fourth, there is no baseline unit of measurement used in Schachter and Rutherford. That is, the number of structures per composition are compared, but they are quite probably not comparable because the length of compositions varies greatly (from this researcher's experience). Thus, the number of words **or,** better yet, sentences is not considered in their analysis, and consequently, the results are not particularly meaningful in a quantitative sense.

75

Schachter and Bart (1979) performed a much cleaner analysis, by comparison, of written production across the same Ll groups and across three defined proficiency levels, because they reported relative frequency of errors, and included averages of total number of words produced at each proficiency level. Rutherford[a]s study was more precise in terms of having baseline measures, since we know that in the pilot study (but not the study leading up to the pilot study) he examined only six sentences per subject.

Fifth is the consideration of validity. Schachter and Rutherford examined written compositions of students based on assigned topics, and this is, we believe, a valid method of eliciting learners[a] (quasi-)spontaneous production (issues of monitoring aside). However, Rutherford's study has two major problems in terms of validity: (1) He had students perform a "simulated[W] composition writing task, for which learners were to provide a topic sentence for each of six assigned titles. The procedure and purpose of this task, Rutherford admits, were difficult to convey to learners. (2) Related to this "simulated composition task[n] are the problems with analyzing a sentence produced out of context (i.e., independent of other sentences in a text with the same discourse topic), for the purpose of ascertaining discourse-syntactic transfer from the Ll and thereby characterizing J and M IL development. What can in fact be generalized is only that learners producing individual sentences on a simulated composition writing task tend to construct topic sentences of the said kind. Note that Schachter and Hart (1979) went so far as to disregard the first sentence in their compositions, probably due to the anticipated effect of modelling or even scaffolding that the title has on the topic sentence, especially in learners[a] production.

Sixth, no inter-rater reliability was reported for the coding of structures in either study, and since constructs such as TCs are inferential categories, this is yet another methodological weakness.

Lastly, the statistical analysis of the data in the previous studies, if performed at all, does not point convincingly to the significance of all of the claimed differences in IL production between J and M.

3.4.2 Sampling of Subjects

Writing samples were drawn from backfiles of English Language Institute (ELI) placement tests at the University of Hawaii at Manoa, and selection was based on the following criteria: subjects had to (1) be native speakers of J or M only (no other Chinese dialects); (2) have recorded TOEFL scores between 450 and 599; (3) have taken the ELI placement examination between 1980 and 1985; (4) have lived in the United States for less than three years. From the population defined above, 35 compositions were randomly sampled for each range of TOEFL scores; i.e., 35 J with TOEFL scores 450–499, 35 with TOEFL 500–549, and 35 with TOEFL 550–599. The same procedure was followed for M subjects. The design is shown in Table 3.1 below, and in Table 3.2 and Table 3.3 the results of sampling are shown in terms of mean TOEFL scores per cell and mean length of residence in months, respectively.

Table 3.1

Subjects

Proficiency level
(by TOEFL scores)

L1	450–499	500–549	550–599
Japanese	n= 35	35	35
Mandarin	35	35	35

Table 3.2

Means of TOEFL Scores By Group

Proficiency level
(TOEFL scores)

L1	(450–499)	(500–549)	(550–599)
Japanese	476.91	517.91	576.66
Mandarin	473.43	521.71	570.43

Table 3.3

Mean Length of Residence (Months) by Group

(TOEFL scores)

L1	Level 1	Level 2	Level 3
Japanese	12.51	8.49	8.34
Mandarin	26.26	8.51	10.86

Unfortunately, the nature of the population at the English Language Institute (ELI) does not yield enough subjects for a separate category of even lower level learners (e.g., TOEFL 400-449); therefore, we cannot make generalizations for that level of J and M learners (as e.g., Schachter and Hart 1979 did).

The baseline unit of analysis was established as 12 T-units (see definition in Appendix A). This text size was determined by tallying the length of each composition available for sampling. It was apparent that many compositions had well over 30 T-units, whereas others had only 10 or so T-units; in order to have a satisfactory n-size to work with (e.g., 30-40 subjects per cell in Table 3.1), 12 T-units represented the largest amount of text that could be analyzed across most compositions. Choosing 15 T-units, for example, would have reduced the n-size to less than 30 subjects per cell.

Therefore, all compositions were first divided into T-units. Second, sampling of the 12-T-unit passages from each composition was done with a table of random numbers to determine the initial T-unit from which 12 consecutive T-units (and any non-syntactic units contained within 12 T-units) would be sampled. The writing samples were then photocopied, and a coding sheet with all biodata was kept separate.

3.4.3 Task

The data base for the present analysis is the writing samples for ELI placement testing purposes between 1980-1985. Students taking the test are instructed to choose one of four given topics and to write on that topic during a maximum time of 50 minutes. Topics include "Describe

the ideal friend[W] "What should the role of a husband or wife be?", "How my family and I celebrate our favorite holiday," and so on. These titles and the task itself are comparable to those used by Schachter and Hart (1979), Schachter and Rutherford (1979), *Rutherford* (1983), and Rutherford and Altman (1985).

A total of more than forty composition topics were written about, as the choice of possible topics changes from term to term at the ELI. Generally speaking, most students write only one draft, making few corrections as they proceed, and the product is on average one to two pages in length (single-spaced), or approximately 25 T-units (see Measures, below).

3.4.4 Design

The basic design of this study is a 2-way analysis of variance in which the independent variables are (1) Ll background, either Japanese or Mandarin, and (2) level of ESL proficiency, as determined by TOEE'L scores: low (TOEFL 450-499), medium (500-549), or high (550-599). Control variables are (1) length of residence, which is less than three years; (2) task, which is the (more or less) sponta- neous writing of compositions in a timed ,placement test; (3) the period of admission into the program at the English Language Institute at the University of Hawaii, namely, 1980-1985; and (4) roughly, academic background.

The dependent variable is simply the syntacticization of Subjects from Topics. The measures of this variable appear in the following section, and they are taken mostly from Rutherford (1983). The ordering of measures is intended to show rather crudely the diachronic development that takes place, but this is only within the numerals (2) and (7) (e.g., a -> b -> c), not across numerals. The first measure, T-units, is not a dependent measure, but rather, a baseline unit upon which to make cross-group comparisons.

In addition to the independent and dependent variables identified above, there are several possible moderating and intervening variables. First of all, composition topics (i.e., assigned titles) are a potential source of variation in the data, but due to the constraints imposed by "convenience sampling[n]" in order to meet the other criteria (independent and control variables), it was not possible to limit the analysis to a fixed topic. In the end, however, there is a randomizing effect on topic

due to the number (40) and sampling of possible topics, and thus we can assume a negligible effect for this variable.

Secondly, even within the range of length of residence, there might conceivably be differences in the production of learners who have just arrived from their home country, and those who have already resided in an ESL community for up to three years. It is to be noted that the sampling procedure of subjects unfortunately resulted in the group of lowest level M learners having a longer length of residence (approximately double) than that of J at the same level, and longer than any other sub-group in the sample. Due to the grouping on the basis of standardized test scores, however, this difference is not expected to bear on the results to any serious extent.

Intervening variables are related to such unknown and undeterminable factors as: degree of monitoring of task, influenced by fatigue; transfer of training; the effect of the testing session itself on learners' performance; and the effect of the amount of explicit instruction in English expository writing and grammar.

3.4.5 Measures

Measures used in the study follow, along with examples; due to space limitations, an operationalizable description of all the criteria is not provided here, although a more complete description appears in Appendix A. For ease of reference in this section and in the following chapters, "2a, 2b, and 2c" will refer to the three stages said to be representative of J learners[1] development (see Table 2.2); by contrast, the labels "7a-7f" will refer to the stages claimed for M learners[1] IL development (see Table 2.1). These coding conventions are consistent with those in Appendix A, where they are explained more fully.

(1) T-units: T-units (Hunt 1977) were coded as they are conventionally coded in the ESL literature (cf. Larsen-Freeman 1983). Subordinate or adjunctive clauses were included in the same T-unit as the main clause, except when the clauses were bounded by terminal punctuation. Sentences without verbs, which are not generally included in T-unit analyses, were coded as "non-syntactic units," and these were important in this analysis since the absence of a verb

in the predicate signals in many cases the most primitive of TC constructions.

Schachter and Hart (1979) argued that T-units were not appropriate baseline measures for IL data, but since the T-unit is based on the existence of a main verb, this is actually the most suitable index for comparing the occurrence and use of Topics, Subjects, and PRO-drop.

In China, many kind of TV program
showing / and the children watching them./
Because can learn many thing from TV./
(3 T-units)

(2a) Sentence-initial locative + VP

Japan, got lot of problem nowadays.

(2b) Sentence-initial locative with in + Sentence

In Japan; wife is expected to stay at home to take care of children.

(2c) Existential there + VP + locative with in

There were always many people living in the small town surrounded by mountains and had no way to communicate with others.

(3) Sentence-initial if-constructions

If you can teach a student in their preference way; I think that's the best way to teach.

(4) Sentence-initial when-constructions

When the people is comfortable in the situation, he or she is not unhappy at least.

(5) Extraposition with sentence-initial it

It is necessary to treat others as he hiself is treated.

(6) Passives

He was decided to be killed.

81

(7a) Topic-Comment: NP (TOP) -∅ (SUBJ)- VP

Women go out of home and look for **a** job
are usual today.

(7b) Topic-Comment: NP (TOP) -NP **(SUBJ)** VP

Older person, Katsu changed my life
significantly.

(7c) <u>There</u>-VP-VP

Therefore we could say that there have good
families got alot of children.

(7d) <u>There</u>-VP-to VP

There is already enough people to populate
earth.

(7e) <u>There-VP-Relative</u> clause

There are **alot** of people who thinks that way.

(7f) Subject-Predicate **(Infinitive/Gerund** as Subject)

To stick to those important rules isn't a
hard job for me.

(7') Miscellaneous existential <u>there</u>

There are alot of people with that idea.

(8) Subject-verb agreement

My father go to work six **o'clock (-)**
My father goes to work six o'clock (+)
My two uncle have houses in Hong Kong (IL)

(9) PRO-drop

He said **(∅)** should go to the movie together.

(10) Serial verbs

The wife have to stay home take care **of** kits.

(11) Topic chains

(see example in Section 2.2.5.)

3.4.6 Coding

Inter-Rater Reliability

Two independent, linguistically sophisticated raters were trained according to the coding conventions and instructions established by this researcher (based primarily on Rutherford's 1983 examples), which are found in Appendix A- After the coding procedures were explained to the raters, they practiced coding four texts of approximately 12 T-units each. Then, after further discussion of procedures, raters independently coded a total of 60 additional texts, or 25% of the entire corpus (each text approximately 12 T-units in length), which had been randomly selected from the corpus according to three parameters: L1, proficiency level, and place in composition from which the texts were extracted (i.e., the number of the initial T-unit from which the 12-T-unit text began). Thus, the texts rated by this researcher and two others could be considered for all intents and purposes representative of the corpus.

Inter-rater reliability scores were calculated on the basis of number of instances of agreement divided by the total number of cases of agreement plus disagreement for each measure (and sub-measure) and for each text. The criterion rating was that of this researcher (Rater 1), thus the purpose of the procedure was to ensure that the categories were indeed "psychologically real" or rateable, and secondly, that this researcher's coding was reliable (i.e., consistent). Thus the calculations were made for Rater 2 relative to the criterion, and for Rater 3 relative to the criterion. Although it is undisputable that Cohen's kappa coefficient provides a more conservative estimate of agreement which can be used in studies such as this, due to the number of measures and texts involved, percentage agreement only was calculated here (Cohen 1960).

This procedure met with mixed results, due it appears (perhaps ironically so) to the processing demands imposed by the task itself. The first set of scores represents the agreement of the identification of boundaries (i.e., T-units) and structures to be coded (e.g., existential there). In this area, all but extraposition and passives received a fairly high level of identification, as Table 3.4 indicates. Given the explicit linguistic description of most of these structures (with those in 7a and 7b considered the most inferential), and the relative success with which raters identified tokens in practice sessions (using a much smaller corpus), it appeared that the task of

reading through unflagged texts to identify potentially codeable structures resulted in some structures going undetected; this was especially true in the case of 'dummy elements" or redundant elements which are usually highly predictable and thus easily missed.

With this understanding, and given constraints on time and resources, it was necessary to calculate the level of agreement on the identified features only, with no further coding of texts. Thus, for extraposition and passives, whose level of identification was initially low, it must be assumed that the difficulty in coding arose from the process of coding and not because the categories are inferential (see descriptions in Appendix A).

Table 3.4

Inter-Rater Reliability

Percentage Agreement on Identification
and Coding of Features

Measure		Rater 2 Identif.	Rater 3 Identif.	% AGR on Coded Features (R2+R3)
T-units		96%	95%	> 90%
S-V AGR	(+/-)	89%	93%	> 90%
If/When		88%	88%	> 90%
Locatives	(2s)	77%	77%	> 90%
TCs/Exist	(7s)	80%	73%	> 90%
Extrapos/Pass.		<SO%	<SO%	> 90%

Of the structures actually identified by raters, the level of agreement was, however, quite satisfactory, ranging from upwards of 90%, and well above 95% in most cases. This score was obtained for both raters on all measures, including the sub-categorization of measures with multiple levels. Due to the low frequency of PRO-drop and serial verbs in the texts sampled, they were not, as it turned out, included in the inter-rater reliability scores.

The coding of the remainder of the data was carried out by this researcher only, according to the guidelines in Appendix A.

84

3.4.7 Analysis

The research design warranted a 2-way analysis of variance (ANOVA) to examine differences between and within groups on each of the dependent measures by Ll (J or M) and by proficiency level (low, medium, or high). In addition, ANOVAs of several computed measures were also performed.

To test Hypothesis (1), which compared J and M production of TCs, an ANOVA was performed with the dependdent measures of 7a (NP-∅-VP) and 7b (NP-NP-VP). Another related category was also included here, namely topicalized prepositional phrases (i.e., PP-NP-VP), which were coded "7bb" (as in Appendix A).

Hypothesis (2) compared J and M production of heavy Subjects; this was tested with a qualitative comparison of J and M constructions in 7a (NP-∅-VP). Hypothesis (3) predicted differences in J and M production of existentials with there. This was tested with an ANOVA of total. existentials produced with there, which was a computed measure based on the sum of existentials in 7c, 7d, 7e, 7' (miscellaneous there constructions; see Appendix A), and 2cs, which were double-coded anyway as sentence-final locatives and existentials.

Hypothesis (4) examined the production of extraposition, and this was tested with an ANOVA of that dependent measure.

Hypothesis (5) predicted cross-linguistic differences in total production of dummy Subjects; this was addressed by computing total scores of the dependent measures of extraposition with it, plus the computed score of total there constructions used to test Hypothesis (3). Note, however, that for purposes of simplicity this analysis did not take into account cases other than extraposition in which the pronoun it served as a dummy element.

Hypothesis (6) compared J and M production of preposed locatives, and this was tested with an ANOVA of the dependent measures 2a and 2b (i.e., loc NP-∅-VP and loc NP-NP-VP).

Hypothesis (7) was tested by an ANOVA of the dependent measure of passives.

To test Hypotheses (8) and (9), respectively, ANOVAs of the dependent measures of sentence-initial if and when clauses, separately and then combined, were performed.

Lastly, and perhaps most importantly, for Hypothesis (10), an ANOVA was used to test. the syntacticization of structures represented in Rutherford's model (namely, 2a > 2b > 2c; and 7a > 7b > 7c > 7d > 7e > 7f), and also, those other dependent measures whose production was predicted to correlate with language proficiency (e.g., agreement, if/when constructions, dummy Subjects, and passives). A crude form of implicational scaling was also performed by plotting the means of the same sets of dependent measures with barcharts comparing production on the basis of proficiency level.

The acceptable level of probability was set at p < .05 for all hypotheses.

CHAPTER IV

RESULTS

The frequency statistics for the analyses, which include the sums, means, and standard deviations of the scores for each dependent measure, broken down by L1 and proficiency level, are displayed in Table A in Appendix B. Figure 4.1 consists of a series of barcharts showing relative frequency of production of the various dependent measures in Hypotheses (1) to (9), which are classified according to means for J versus M. Figures 4.2 to 4.11 correspond to the results for Hypothesis (10), which deals with differences across proficiency levels on a number of dependent measures. Figures 4.12 and 4.13 display differences in J and M production of PRO-drop and serial verbs, which are not directly addressed in Hypotheses (1) to (10).

Below, the results are reported in terms of the ten original hypotheses.

4.1 Hypothesis (1) M learners produce more TC constructions than J.

Frequencies of TCs in 7a and 7b (see Table A, Figure 4.1) reveal that M produced almost double the number of 7a TCs (NP-∅-VP) as J, with group sums of 21 versus 11, and means of 0.20 and 0.10, respectively. However, based on the ANOVA for this measure, this difference is not statistically significant (F=2.22, df=1/204, ns). Similarly, by total counts, M produced more of TC 7b (NP-NP-VP) than J, with sums of 35 versus 28, and means of 0.33 and 0.27, respectively, but this difference was not significant (F=0.66, df=1/204, ns). For TC 7bb (PP-NP-VP) also, sums and means were M 26 and 0.25, J 14 and 0.13, but again the difference was not significant (F=2.56, df=1/204, ns), hence the hypothesis was not supported.

4.2 Hypothesis (2): Among the TCs referred to in (1), M also produce more sentential or "heavy Subjects" than J.

The results of this analysis were that there were relatively few' cases of sentential Subjects of the type discussed by Rutherford (1983): 7 for M, and 2 for J, with corresponding means of 0.06 and 0.02, respectively. Thus there was a clear trend for more of these constructions to be produced by M than J, as hypothesized, but an ANOVA on

87

Figure 4.1

**Summary of L1 Group Comparisons
in Hypotheses (1) to (9)
Classified by Dependent Measures**

88
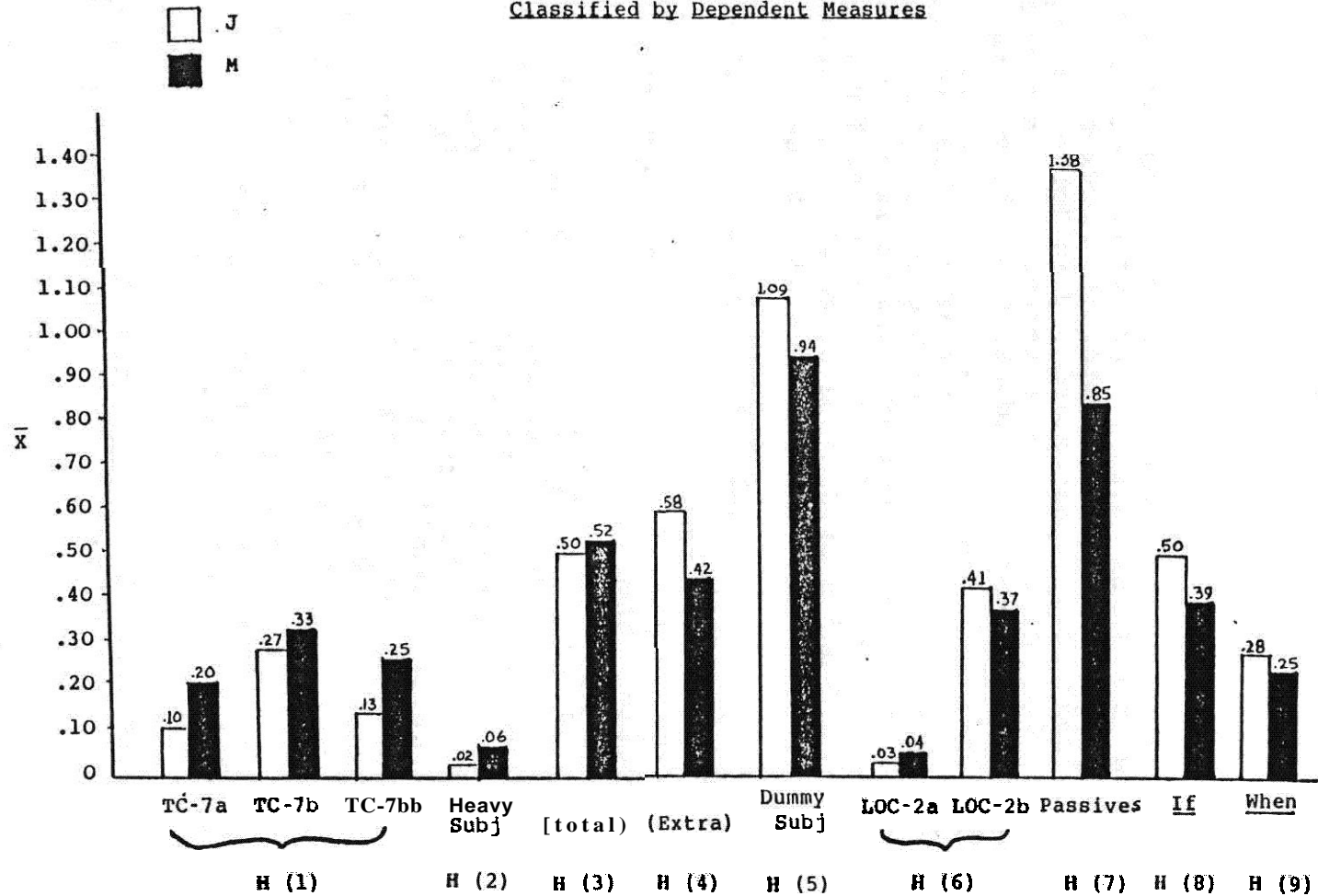
this test was not performed, due to the low production frequencies.

4.3   Hypothesis **(3)**:   M produce more existential constructions than J, marked with the non-referential element <u>there</u>.

The results of this test were that M and J produced almost exactly the **same** number of existential constructions with the expletive <u>there</u>. Group sums for M and J, were 55 and 53, respectively, and mean scores were 0.52 and **0.50.** The differences were, therefore, not significant **(F=0.03, df=1/204, ns)**, thus this hypothesis was not supported. **Subclassification** of the different types of <u>there</u> constructions (see Table A) revealed that M tended to produce more of **7c** and **7d**, and J more of **7e** and 7', but none of the trends were significant.

4.4   Hypothesis **(4)**:   **J** produce more extraposition constructions (marked with non-referential <u>it</u>) than M.

By total frequencies, J produced more extraposition than M (sums were 61 and 44, and means were **0.58** and 0.42, respectively), but this difference was not significant **(F=2.53, df=1/204, ns)**. This hypothesis was not supported.

4.5   Hypothesis **(5)**:   The combined total of dummy Subjects (**it** and <u>there</u>) is **greater** for J than M.

J produced on the whole more expletives than M (group sums were 114 and 99, and means were **1.09** and **0.94,** respectively), but again, this difference was not significant **(F=2.53, df=1/204, ns)**; therefore, the hypothesis was not supported.

4.6   Hypothesis **(6)**:   J produce more sentence-initial locative constructions, with or without the existential <u>there</u>, than M.

This hypothesis was tested with **ANOVAs** of 2a and 2b scores.   The results showed that for **2a**, the group sums were very low: 3 for J and 4 for **M**, means 0.03 and 0.04, respectively, and this difference was not significant **(F=0.11, df=1/204, ns)**. For **2b**, although the total frequencies were greater, 43 for J and 39 for M, with mean scores of 0.41 and 0.37, respectively, again the difference between the two groups was not significant **(F=0.15, df=1/204, ns)**. Even including scores for **2c**, as a score for total locatives in **Rutherford's** model (Table **2.2**), the sums were 57 for J and 54 for M (means 0.54 and 0.51,

respectively), and this was not a significant difference across Ll groups (F=0.15, df=1/204, ns). In short, Hypothesis (6) was not supported.

4.1 Hypothesis (7): J produce more "true" passives than M.

The results of this test revealed that J produced a total of 145 passives, compared with 89 for M (means were 1.38 and 0.85, respectively), and this difference was highly significant (F=8.95, df=1/204, p < .005) as shown in Table 4.1 below. Thus this hypothesis was supported with a significant main effect for Ll in passive production. There was also a significant effect for proficiency and a 2-way interaction of passive with proficiency, but this will be reported under the heading of Hypothesis (10).

Table 4.1

ANOVA of the Dependent Measure of passives by the Independent Variables of Ll and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance	
Main Effects	29.11	3	9.70	5.81	0.00	****
Ll	14.93	1	14.93	8.95	0.00	***
PROF	14.17	2	7.09	4.25	0.02	*
2-Way Interactions	13.75	2	6.88	4.12	0.02	*
Ll x PROF	13.75	2	6.88	4.12	0.02	*
Explained	42.86	5	8.57	5.14	0.00	****
Residual	340.40	204	1.67			
Total	383.26	209	1.83			

```
*    p < 0.05
**   p < 0.01
***  p < 0.005
**** p < 0.001
```

4.8   Hypothesis (8):   J produce more sentence-initial _if_ constructions than M.

Total frequencies for the production of preposed _if_ clauses were 52 for J and 41 for M, with mean scores of 0.50 and 0.39, respectively. This difference was not significant (F=0.96, df=1/204, ns); therefore, the hypothesis was not supported.

4.9   Hypothesis (9):   J produce more **when** constructions at the beginning of sentences than M.

The total scores for this measure were very close across groups, J 29 and M 26, with corresponding means of 0.28 and 0.25.   This difference was not significant (F=0.14, df=1/204, ns), thus the hypothesis 'was not supported.   Similarly, the combined total of scores for _if_ and **when** from Hypotheses (8) and (9) revealed that while. J tended to produce more constructions of the said type than M (sums were 81 and 67, and means 0.77 and 0.64, for J and M respectively),   again the difference was not significant (F=1.11, df=1/204, ns).

4.10   Hypothesis (10): There are differences in the effects of **Tp, e.g.,** the amount of **TCs,** based on L2 proficiency.   Thus,. as learners syntacticize their constructions, they consequently produce fewer constructions at the lowest levels of Rutherford's model, fewer other constructions with the same discourse function (e.g., conditionals), and more Subject-verb agreement.

The discussion of results for this hypothesis will be ordered according to the list of measures in Appendix A and Appendix B (Table A), beginning with the **syntacticization** of locatives (see **Rutherford's** prediction in Table 2.2). Since **L1** was not found to have a significant effect in most of the hypotheses above, in presenting the results for Hypothesis (10), the scores of J and M groups are pooled unless indicated otherwise.

4.10.1   Locatives (2a > 2b > 2c)

It was predicted that the development **of** existentials **out** of locatives, claimed to be characteristic of J learners' **ESL,** would (roughly) follow a three-step progression: (i) locative NP **+** predicate (2a); (ii) locative prepositional phrase with Subject NP or **non-**referential pronoun _there_ (2b); and lastly, (iii) existential sentence introduced by _there_ with a **sentence-**final locative prepositional phrase (2c). This

developmental trend was predicted to be more typical of J data than M (see Hypothesis 6), but also, it was expected to correlate with the proficiency level of learners in a cross-sectional design. Proficiency levels will be referred to as low (1), medium (2), and high (3) in the discussion which follows.

In Figure 4.2, the results of production frequencies are charted out along the horizontal axis, which is intended to represent the hypothesized developmental order from left to right, as a kind of implicational scale. (Note that 2bb is also included, namely In + temporal + NP/there VP, since it figured prominently in the analyses of Curran 1984 and Duff 1984).

The interpretation of this figure is that at the lowest developmental stage, the 2a structures are produced by level 1 and 2 learners (although the means are very low for both), and they are not at all produced by level 3 learners, as predicted. This finding is not significant, however (F=1.48, df=2/204, ns), due probably to the low production frequency.

For 2b, the cross-proficiency production pattern depicted in Figure 4.2 for 2a is reversed (although the low frequency of production in 2a must again be taken into consideration), with levels 1 and 3 displaying higher mean scores than level 2 (identical for the computed total of preposed adverbials 2b and 2bb, moreover). The difference across proficiency levels is not significant (F=1.00, df=2/204, ns), and neither is the scaling pointed in the correct direction, with, say, a higher mean for level 2 than levels 3 or 1.

Lastly along the developmental scale, for 2c the scaling of means by proficiency levels is opposite that expected, since level 3 learners should produce more 2cs than learners at either level 1 or 2. The differences which do exist are not significant regardless of direction (F=1.68, df=2/204, ns). This prediction is thus not borne out. As a final note concerning the results for locatives, an examination of Table A shows that the breakdown by proficiency for J alone, which might seem to be a more sensible analysis of the locative data, does not provide any more support for the hypothesis than the combined J and M totals.
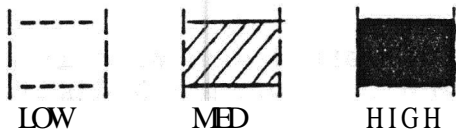
LOW (level **1**) = TOEFL 450–499
MED (level **2**) = TOEFL 500–549
HIGH (level **3**) = TOEFL 550–599

Figure 4.2

Means of LOC Scores (2a, 2b, 2c)
Classified by Dependent Measure and Proficiency Level
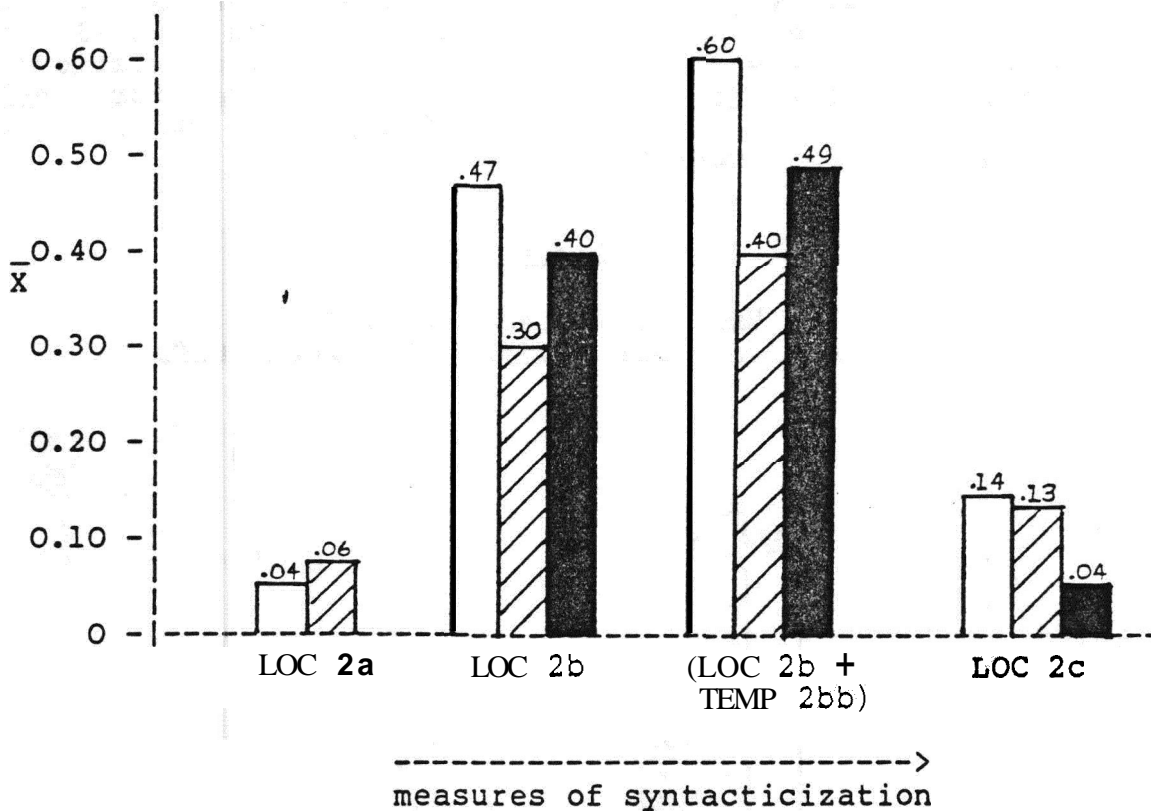
4.10.2 If/When Clauses

Turning next to the results of tests for production of sentence–initial if and when clauses, Figure 4.3 displays the means by proficiency level for each, ,although the two structures are not meant to be linked to each other in any developmental way.   The hypothesized pattern is that lower level learners produce more if and when clauses than more advanced learners, because, it was argued in Chapter II, the structures are used to serve partly pragmatic func- tions, such as topicalization. Notice, however, that this analysis obscures any qualitative differences there might be in the if/when structures produced (any interesting qualitative differences will be reported in the Discussion chapter below). As Figure 4.3 indicates, the general developmental progression within each measure is essen- tially upheld, especially between levels 2 and 3, as the production means decrease inversely with level of proficiency. These trends are not, however, statistically significant (for if, F=1.50, df=2/204, ns; for when, F=0.75, df=2/204, ns). Therefore, the hypothesized difference is not supported.

Figure 4.3

Means of If/When Scores
Classified by Dependent Measure and Proficiency Level
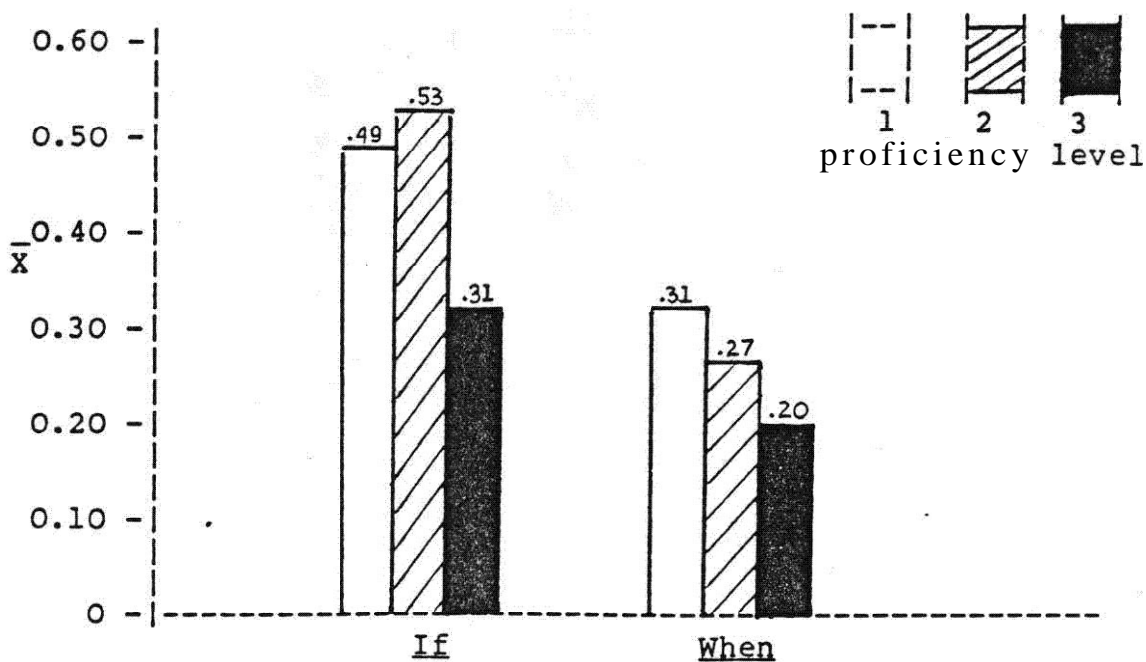
4.10.3 Dummy Subjects

Next, the results for the production of extraposition with _it_, and total production of expletives (_it_ and <u>there</u>) will be reported. The prediction was that there would be, in general, greater production of extraposition and dummy Subjects with increasing level of proficiency, primarily because neither L1 has expletives, but also because they involve a rather sophisticated level and range of English syntax. The IL role of extraposition and dummy Subjects was addressed in previous chapters, though, and there was reason to predict some production at earlier levels, especially in J data. However, production was not anticipated to be greater at lower levels than at higher ones. According to the means plotted in Figure 4.4 (also found in Table A), there is a slight tendency for level 2 to produce (possibly "overproduce") the structures, relative to their counterparts at levels 1 and 3. Furthermore, there is a slight trend for level 3 learners to have higher means than those at level 1. The analyses revealed, however, that there was no significant difference across proficiency levels (for extraposition, $F=0.34$, $df=2/204$, ns; for total dummy Subjects, $F=1.14$, $df=2/204$, ns). Therefore, the hypothesized difference between levels 3 and 1, in particular, was not supported.

Figure 4.4

Beans of Extraposition and Dummy Subject Scores
Classified by Dependent Measure and Proficiency Level
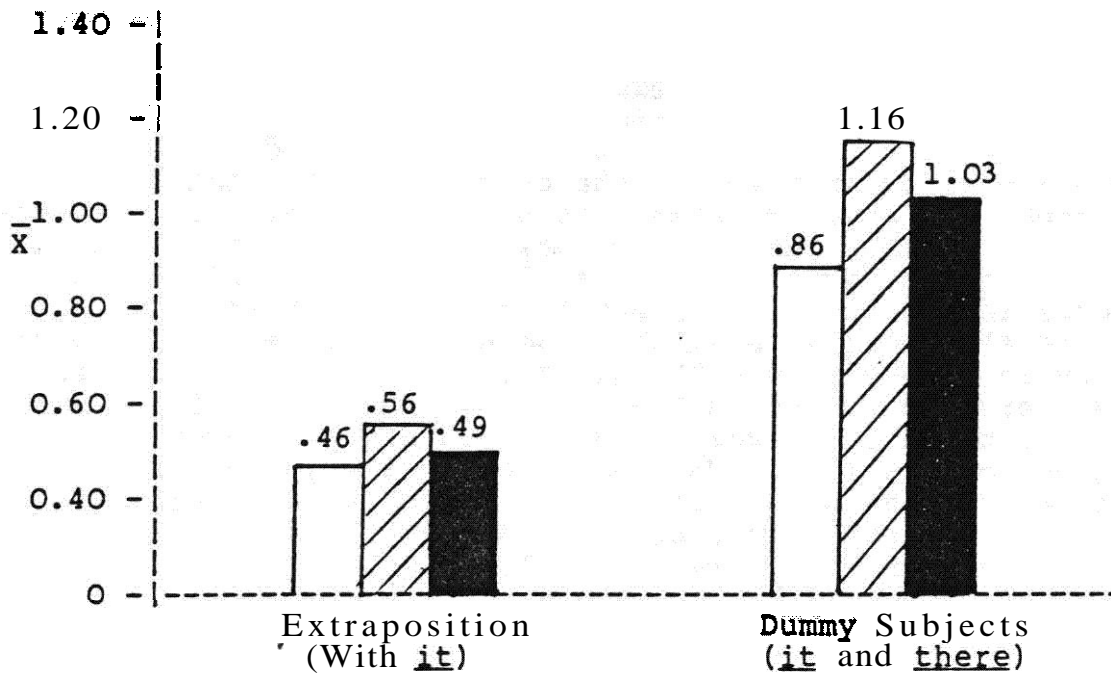
4.10.4   Passives

The mean frequencies for passives are displayed in Figure 4.5, broken down by L1 and proficiency level, due to the interaction effect of these variables.   As with extraposition and overall expletive production, the hypothesized developmental trend was for a general correlation of production with proficiency level, with the possibility for some (IL) overproduction at level 2.   Indeed, the analysis for passives produced the strongest set of results of all the tests.

Figure 4.5

Means of Passive Scores
Classified by L1 and Proficiency Level
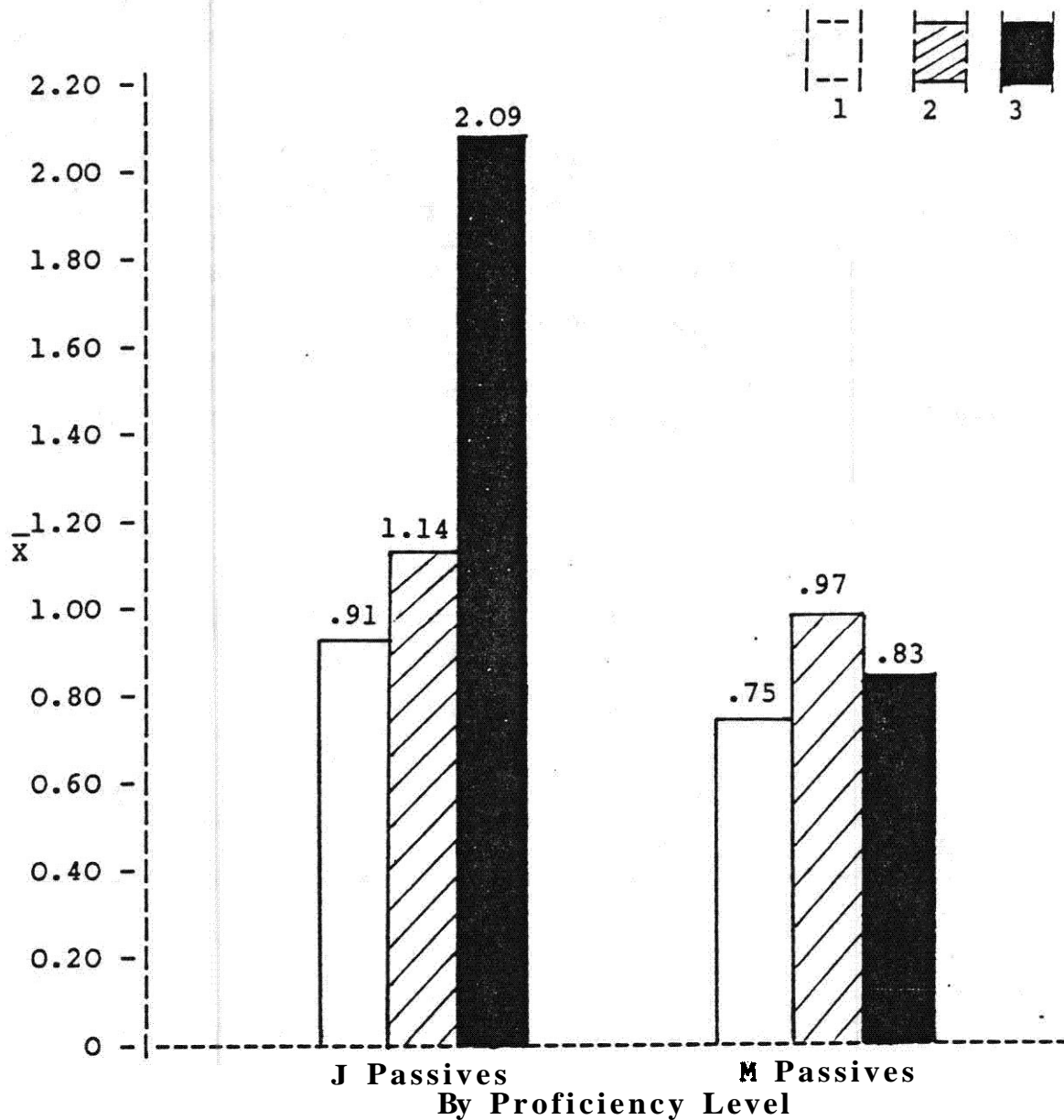
From Figure 4.5, the difference in production across proficiency levels is obvious, and furthermore, it follows the predicted direction. (Note, however, that the breakdown for M in Table A shows that at level 3, they do not produce as many passives as they do at level 2.)

The ANOVA results in Table 4.1 above show that there is a significant main effect for proficiency level ($F=4.25$, $df=2/204$, $p < .05$), and that the 2-way interaction effect of L1 (recall Hypothesis 7) and proficiency level was also significant ($F=4.12$, $df=2/204$, $p < .05$). Thus, of the structures analyzed up to this point in the present study, passives exhibit the strongest differentiation for L1 and proficiency level.

A post hoc multiple-range comparison of the significant finding for passives across proficiency levels was conducted using Tukey's Honest Significant Differences Procedure. This additional test reveals the true source of the significant difference: whether between levels 1 and 2, 2 and 3, or 1 and 3. The Tukey test was performed separately for J and M, due to the interaction effect of L1 and proficiency. The results for J production of passives, shown below in Table 4.2, indicate that the high proficiency group (3) proved to be significantly different from both the low level group (1) and the medium level group (2) at the 0.05 level; but differences between levels 1 and 2 were not significant.

Table 4.2

Tukey's Test of Passive Production Across Proficiency Levels

Japanese Group

Mean	Group (Proficiency)	Low (1)	Medium (2)	High (3)
0.91	Low (1)			
1.14	Med (2)			
2.09	High (3)	*	*	

* pairs of groups significantly different at $p < .05$

The differences in passive production across proficiency levels for M, however, were not significant ($F=0.45$, $df=2/204$, ns). This is shown below in Table 4.3.

98

Table 4.3

Tukey's passive Production
Across proficiency Levels

Mandarin Group

Mean	Group (Proficiency)	Low (1)	Medium (2)	High (3)
0.74	Low (1)			
0.83	Med (2)			
0.97	High (3)			

(no pairs of groups significantly different at p < .05)

4.10.5   TCs and Existentials (7a > 7f)

The next set of structures that were analyzed include those originally discussed by Rutherford (1983) in Table 2.1 (Section 2.3.7.5). The series of initially pre-syntacticized TCs which give way to existentials with there and later syntacticized Subjects, such as geruhds and infinitives, is proposed as a developmental sequence for M in particular. The ANOVAs of production of each of the structures will be reported here, with emphasis on the effect of proficiency. The barchart in Figure 4.6 is intended to show not only differences in production based on relative proficiency for each measure, but also, how the interrelated measures can be scaled (from left to right) from lowest to highest level of proficiency, as an overall index of syntacticization. Also plotted on the chart in Figure 4.6 are the measures of percentage of Subject-verb agreement. Leftmost on the horizontal axis is percentage of total production of finite verbs (present tense) without agreement, which is expected to have higher means for level 1 than level 3. At the opposite end of the scale, i.e., at the pole of optimal syntacticization of Topic/Subject, there is, conversely, a higher percentage of Subject-verb agreement expected for level 3 than for level 1.

99

Figure 4.6

Mean Scores of Existentials with "There"
and Related Structures (TC 7s)
Classified by Dependent Measure and Proficiency Level
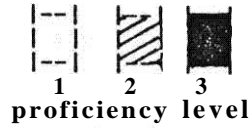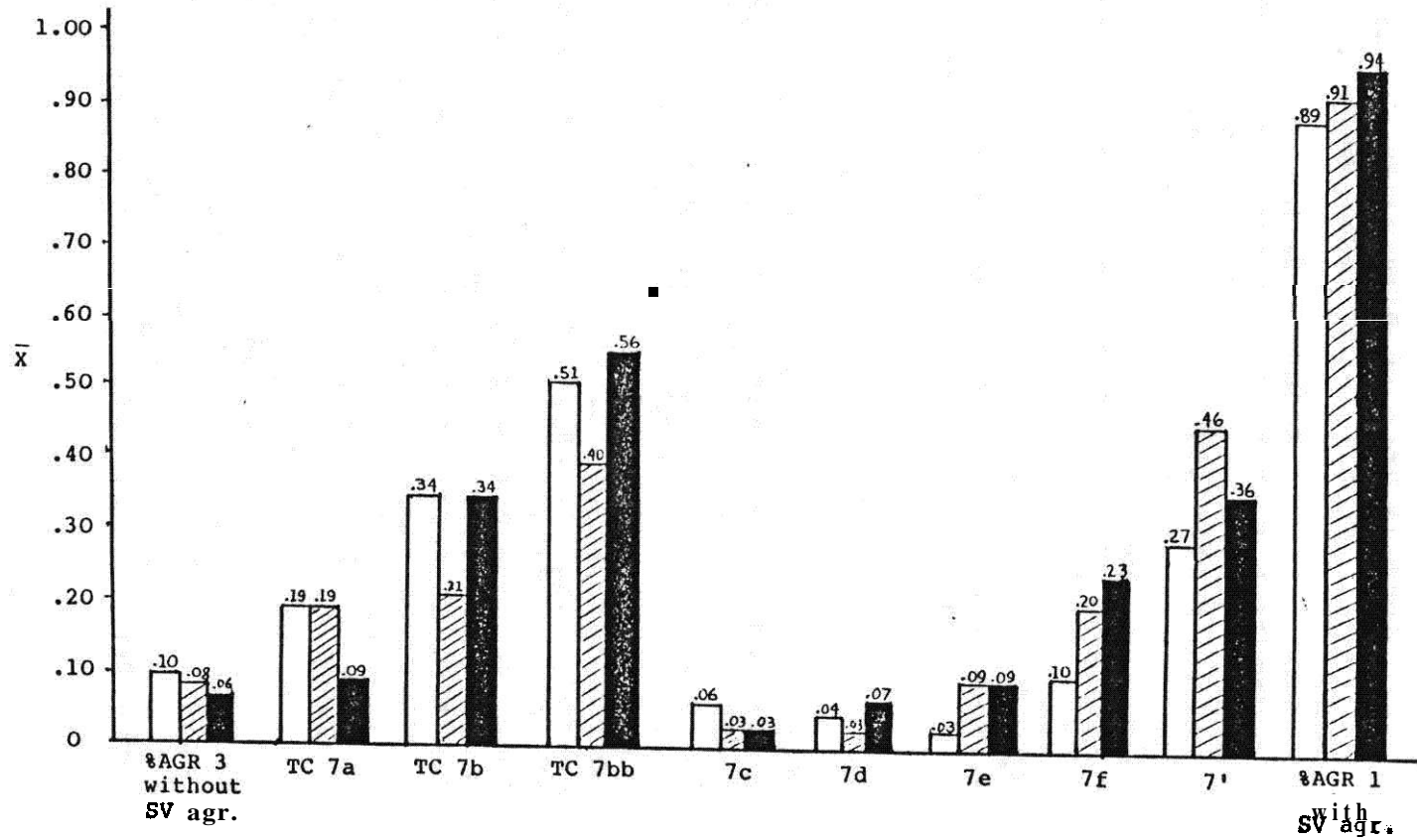
For measures 7a to 7f, according to Rutherford's model, there should be a progressive increase in production for level 3 relative to levels 2 and 1 the further to the right one proceeds along the horizontal (developmental) axis. That is to say that it is predicted that at 7a, level 1 means are higher than those of levels 2 and 3, and conversely, at 7f, the hypothesized trend is for level 3 to produce more structures than levels 2 or 1. This prediction is conceptualized in Figure 4.7 below.

Figure 4.7

Predicted Developmental Pattern for Syntacticization of Topic Based on Rutherford's Model
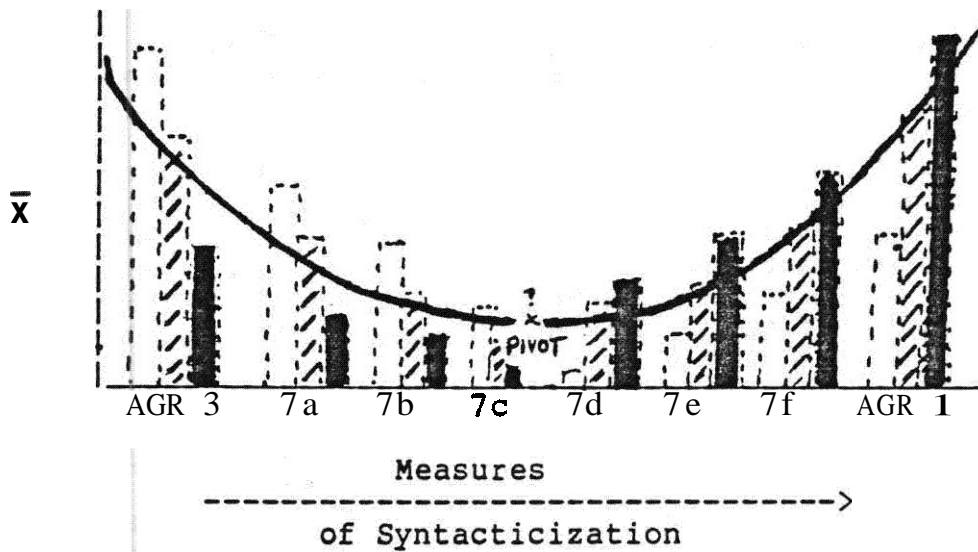
The results as charted out in Figure 4.6, and broken down in Table A (Appendix B), reveal that the hypothesized within-measure differences were, generally speaking, reflected in means for % AGR 3, 7a, 7c, 7e, 7f, and % AGR 1. Those reflecting different trends were 7b (and 7bb), in which there was a persistence of TCs even at level 3 (which, however, was a fairly well motivated and predictable candidate for discourse-syntactic transfer, according to the discussion in Chapter 11); 7d and 7' also exhibited patterns across proficency levels that were slightly different from those expected.

An interesting observation is that between **7c** and 7d there appears to be a shift or pivot from which point higher level proficiency learners begin to produce more of the structures to the right than lower level ones. In other **words,** up until **7c** there is a general tendency for greater production to follow the order of proficiency (level): 1 > 2 > 3. But from **7d** onward? there is a predicted shift in the **order,** which becomes: 3 > 2 > 1.

Another point to be raised in connection with this is that at about this hypothetical pivotal **point,** a construction that Rutherford claimed would not be **produced,** i.e., sentences **like,** "There are many people <u>they</u> work too **hard,**" with the included pronoun <u>**they,**</u> actually occurred in the M data. Caution must be exercised in drawing great inferences from these directional trends **however,** since means for **7c-7e** indicate that very few of these structures were actually produced: a grand total of **8,** **10,** and **14,** **respectively,** out of a total of 2520 T-units **(i.e.,** 210 subjects x **12** T-units **each),** the sum of which represents only 1% of the total corpus.

It is to be noted that Rutherford's model does not have a category **for miscellaneous** existentials with <u>**there,**</u> **such** as **7'** in the present analysis; nor has Rutherford **indicated** that **TCs** at **stage** 7b should be broken down as **NPs** (7b) as distinct from **PPs (7bb),** but this was done to be conservative about the interpretation **of** the Topic "**NP**" at stage **(ii)** in Table **2.1.**

Individual **ANOVAs** for each of the dependent measures plotted in Figure 4.6 did not yield significant results for any of them (disregarding the two measures of Subject-verb agreement for the time being). The only statistically significant result was that M on the whole produced more structures of the kind in 7 than **J,** as determined by an **ANOVA** of the computed sum of 7a through **7f,** which was not addressed in Hypotheses **(1)** through (9) (see Table 4.4; F=4.42, df=1/204, p < .05). A summary **of** the **(non-**significant) results **of** the **ANOVAs** for a main effect of proficiency level on each of these dependent measures follows in Table **4.5.**

Table 4.4

ANOVA of the Total of Dependent Measures 7a to 7f by the Independent Variables of L1 and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance
Main Effects	7.65	3	2.55	1.64	0.18 *
L1	6.88	1	6.88	4.42	0.04
PROF	0.77	2	0.39	0.25	0.78
2-Way Interactions	0.70	2	0.35	0.22	0.80
L1 x PROF	0.70	2	0.35	0.22	0.80
Explained	8.34	5	1.67	1.07	0.38
Residual	317.49	204	1.56		
Total	325.83	209	1.56		

* $p < 0.05$

Table 4.5

Results of ANOVAs for Measures 7a-7'

Measure	F	df	Signif.
7a	1.08	2/204	ns
7b	1.09	2/204	ns
7bb	0.12	2/204	ns
7b+7bb	2.40	2/204	ns
7c	0.41	2/204	ns
7d	0.45	2/204	ns
7e	1.06	2/204	ns
7f	1.44	2/204	ns
7'	1.37	2/204	ns
Total there	1.05	2/204	ns

The hypothesis is, therefore, not supported with a quantitative analysis. Looking across the drafted scale in Figure 4.6, however, there appears to be a gradual tendency toward syntacticization, although there are also some inconsistencies. For example, at the same time that level 3 learners are producing roughly targetlike relative clauses in existentials in 7e (mean=0.09), they produce the same frequency (by mean scores) of 7a's; on the other hand, they produce fewer tokens of the 7a type than level 1 or 2 learners.

4.10.6 Subject-Verb Agreement

Turning to the analysis of Subject-verb agreement, three categories were established: AGR 1 refers to finite verbs in the present tense without agreement; AGR 2 is an intermediate (or IL) category created primarily to account for cases in which the verb was marked for agreement, but the quantified NP did not bear a plural morpheme: e.g., my four brother have; two month are a long time; AGR 3 refers to clear cases of Subject-verb agreement. .The analysis of agreement took two forms. First, based on raw frequencies, ANOVAs were performed, and the results of these are presented in Figure 4.8 according to means compared across proficiency levels (see also Table A, Appendix B for a breakdown of frequency statistics). As was noted above, the trend follows the prediction that the higher the proficiency, the greater the Subject-verb agreement. The opposite is also true; namely, the lower the proficiency, the lower the agreement (or the higher the lack of agreement). The differences charted out in Figure 4.8 are statistically significant for L1, but not proficiency, in terms of overall frequencies for AGR 3 only (lack of agreement). This finding for an L1 main effect is presented in Table 4.6 (F=4.66, df=1/204, p < .05). In Figure 4.9, mean frequencies for AGR 3 are plotted by L1 and proficiency. For AGR 1 and AGR 2 the results were not significant (F=2.86, and F=1.59, respectively, df=2/204, ns).

Table 4.6

ANOVA of the Dependent Measure of Lack of S-V Agreement
(AGR3) by the Independent Variables of L1 and Proficiency

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance	
Main Effects	13.83	3	4.61	3.48	0.02	
L1	6.17	1	6.17	4.66	0.03	*
PROF	7.66	2	3.83	2.89	0.06	
2-Way Interactions	5-49	2	2.74	2.07	0.13	
L1 x PROF	5.49	2	2.74	2.07	0.13	
Explained	19.31	5	3.86	2.92	0.01	
Residual	269.94	204	1.32			
Total	289.26	209	1.38			

* $p < 0.05$

Figure 4.8

Means of Subject-Verb Agreement Scores
Classified by Dependent Measure and Proficiency Level


```
12.00 -|                                                         11.16
                                                       11.14   ┌──┐
11.00 -|                                                ┌──┐    │██│
                                                        │//│    │██│
10.00 -|                                        9.74    │//│    │██│
                                               ┌──┐     │//│    │██│
 9.00 -|                                       │  │     │//│    │██│
                                               │  │     │//│    │██│
 8.00 -|                                       │  │     │//│    │██│

        1.09
 1.00 -| ┌──┐   .94
         │  │  ┌──┐
_0.90 -| │  │  │//│
X        │  │  │//│
 0.80 -| │  │  │//│
         │  │  │//│    .63
 0.70 -| │  │  │//│   ┌──┐
         │  │  │//│   │██│
 0.60 -| │  │  │//│   │██│
         │  │  │//│   │██│
 0.50   │  │  │//│   │██│
         │  │  │//│   │██│
 0.40 -| │  │  │//│   │██│
         │  │  │//│   │██│
 0.30 -| │  │  │//│   │██│           .21
         │  │  │//│   │██│          ┌──┐
 0.20 -| │  │  │//│   │██│          │  │   .13
         │  │  │//│   │██│          │  │  ┌──┐   .09
 0.10 -| │  │  │//│   │██│          │  │  │//│  ┌──┐
         │  │  │//│   │██│          │  │  │//│  │██│
    0 -|─┴──┴──┴──┴───┴──┴──────────┴──┴──┴──┴──┴──┴──

           AGR 3             AGR 2            AGR 1
          without            with IL          with
         agreement         agreement        agreement

         ──────────────────────────────────────────>
              measures of syntacticization
```
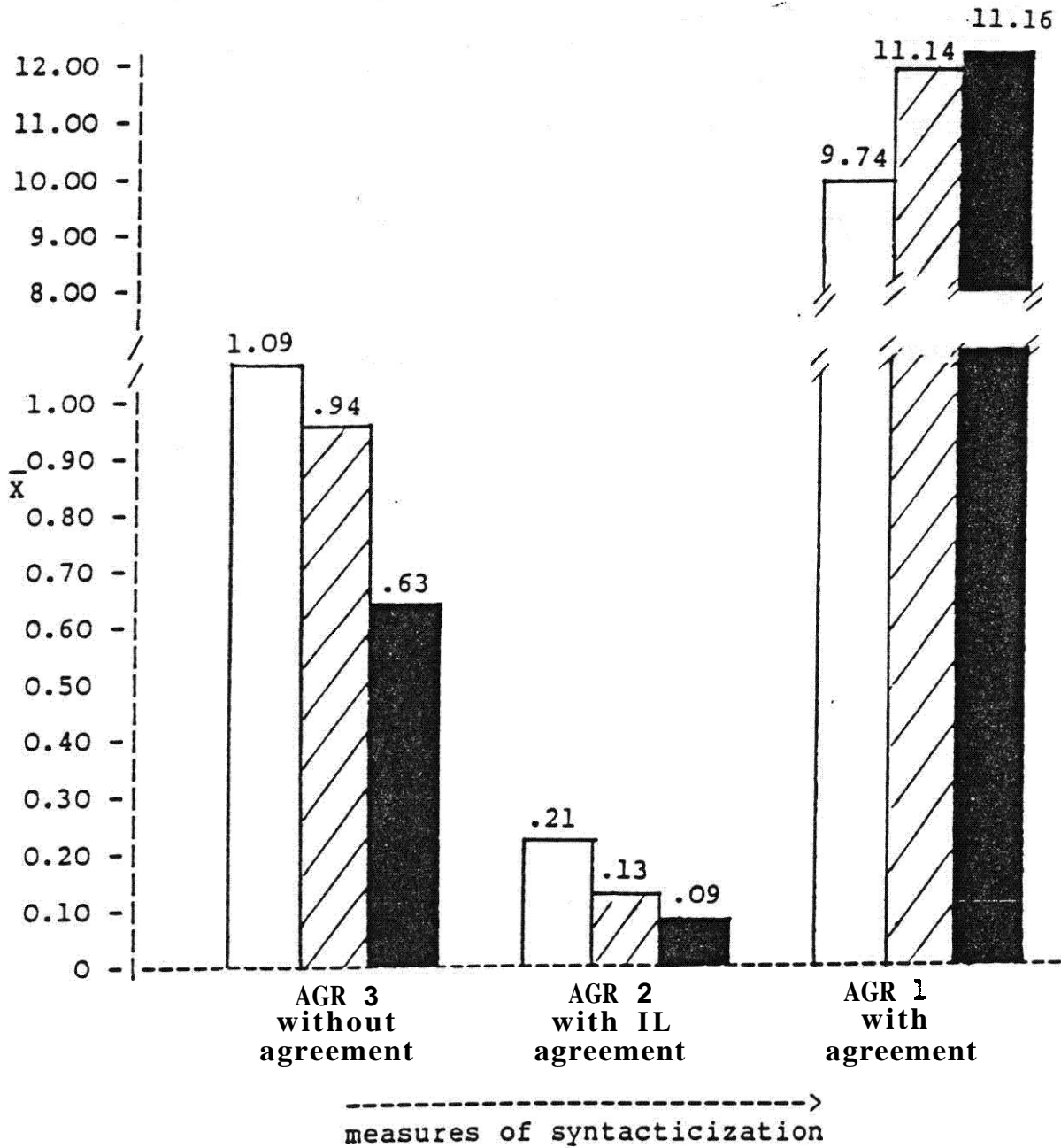
Figure 4.9

Means of (AGR 3) Subject-Verb Agreement Scores
Broken Down by Ll and Proficiency Level
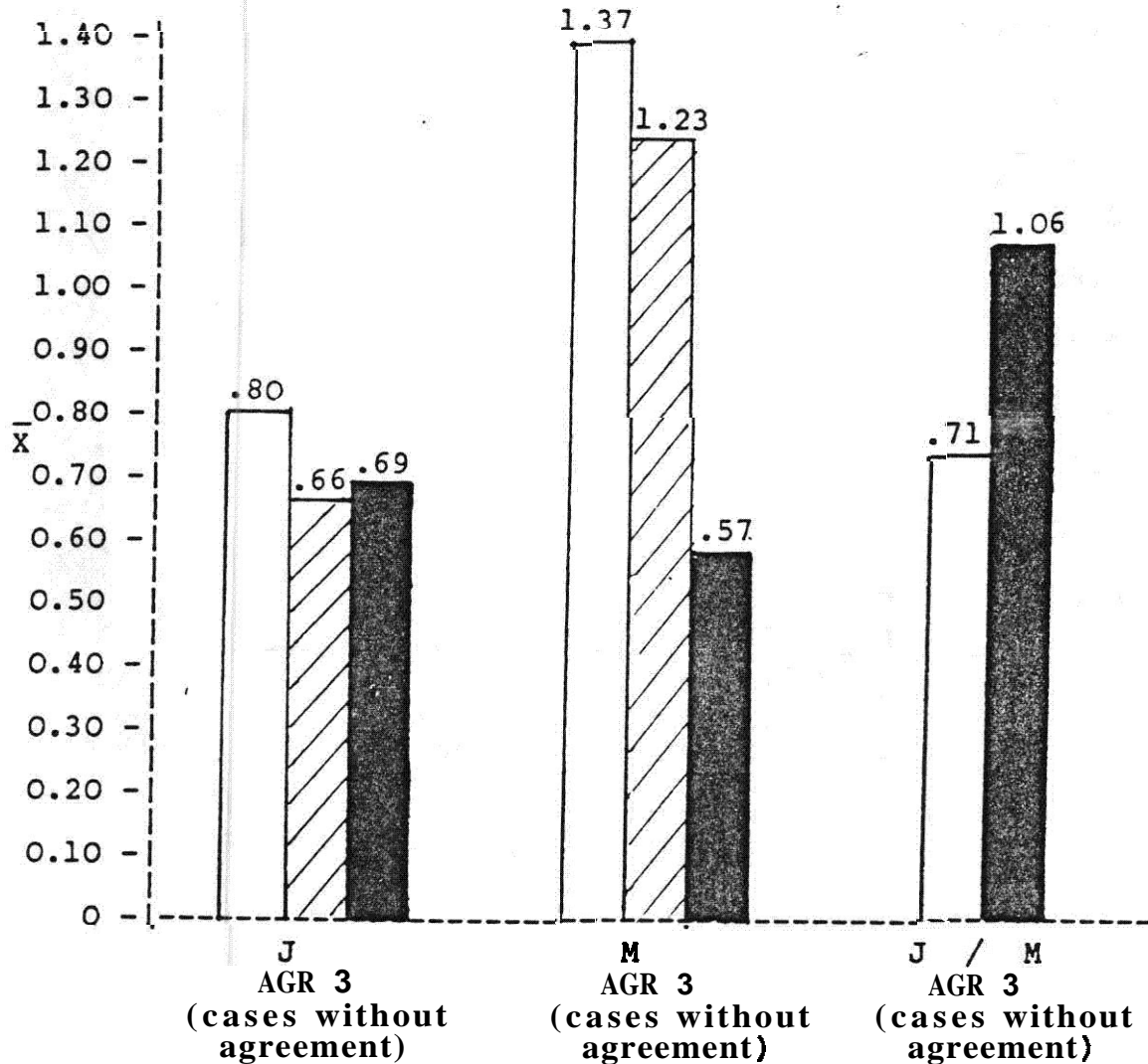
A more meaningful treatment of Subject-verb agreement calculates the rate of agreement (e.g., percentage) in terms of the production of AGR 1, AGR 2, and AGR 3 divided by the sum of possible contexts for agreement. This is shown in Figure 4.10.

Figure 4.10

Means of Percentage Subject-Verb Agreement
Classified by Dependent Measure and Proficiency Level
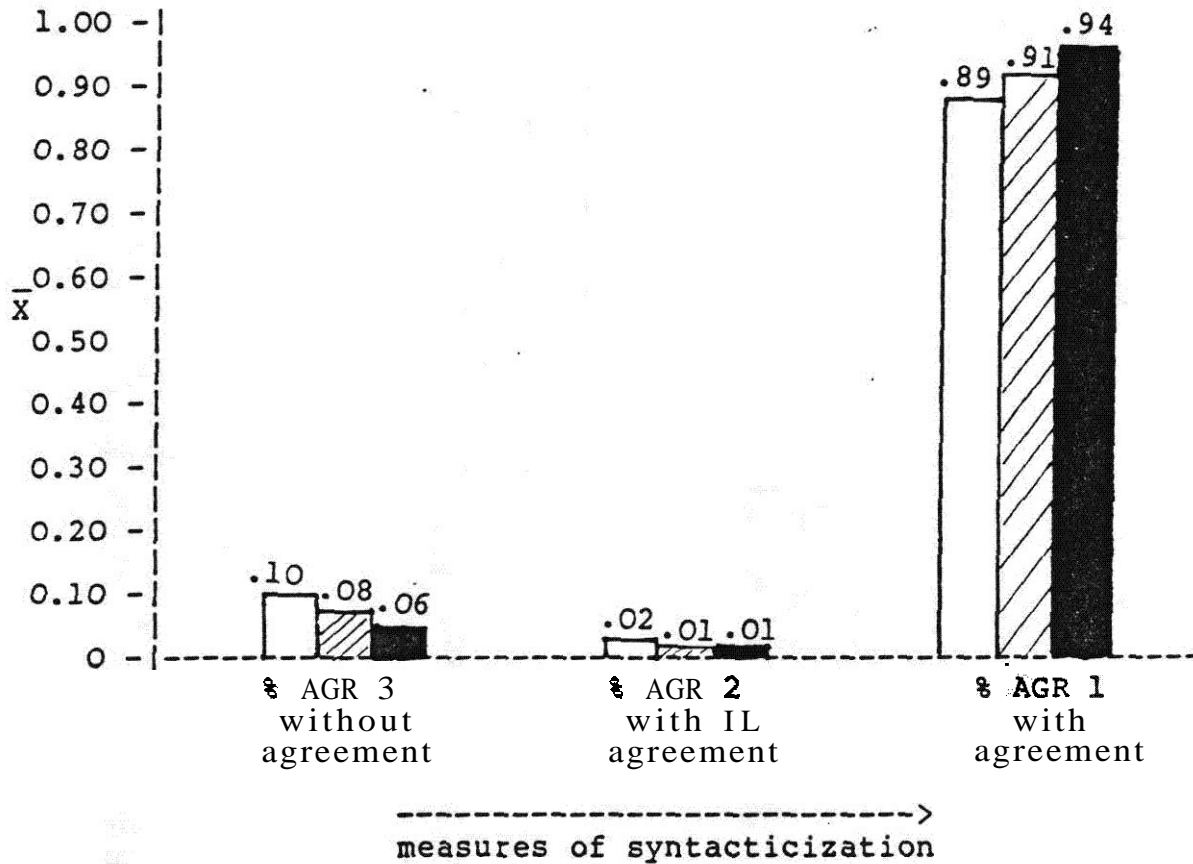
In order to perform ANOVAs on the rate of Subject-verb agreement, the ratios for AGR 1, AGR 2, and AGR 3 were converted by an Arc Sin transformation, and this guarded against analyses based on artificially inflated means. The results of the ANOVAs of agreement revealed that while all groups were producing very high levels of agreement (AGR 1), upwards of 86% (see Table A, Appendix B), there was still a significant main effect for L1, but not for proficiency (for L1, F=4.93, df=1/204, p < .05; for proficiency level, F=2.61, df=1/204, p < .05). The ANOVA results for this measures are presented in Table 4.7, and the L1 difference is shown in Figure 4.11, in terms of a percentage ratio.

108

Table 4.7

ANOVA of the Rate of Subject-Verb Agreement (AGR 1)
by the Indewendent Variables of L1 and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance	
Main Effects	0.96	3	0.32	3.38	0.02	*
L1	0.47	1	0.47	4.93	0.03	*
PROF	0.50	2	0.25	2.61	0.08	
2-Way Interactions	0.13	2	0.06	0.67	0.51	
L1 x PROF	0.13	2	0.06	0.67	0.51	
Explained	1.09	5	0.22	2.30	0.05	*
Residual	19.34	204	0.10			
Total	20.42	209	0.10			

* $p < 0.05$

Figure 4.11

Means of Percentage Subject-Verb Agreement
Classified by Dependent Measure and Ll

measures of syntacticization
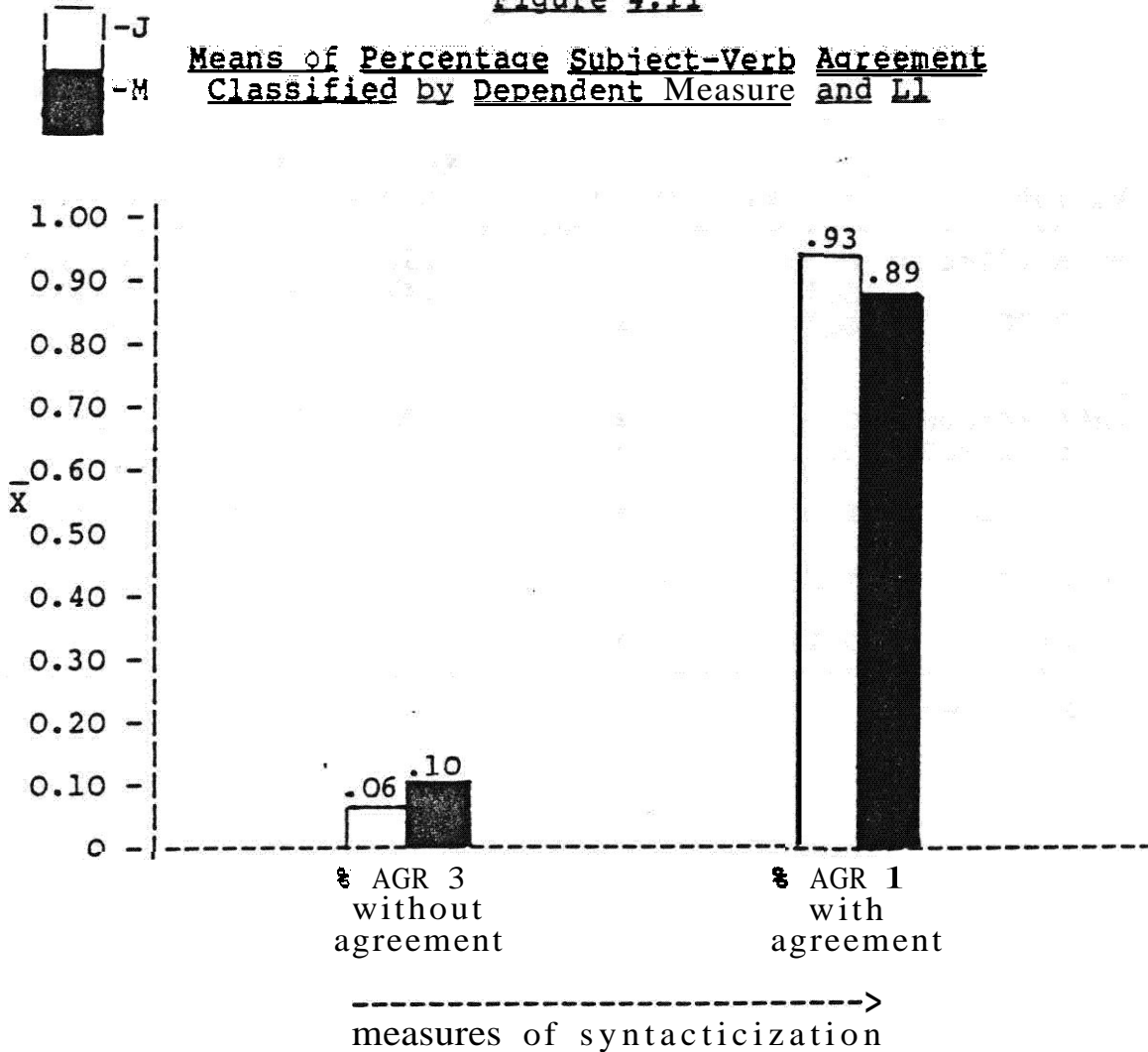
An ANOVA of the rate of lack of agreement also yielded statistically significant results, but again, only for Ll (F=6.96, df=1/204, p < .01; for proficiency, F=2.74, df=2/204, ns), as shown in Table 4.8 and Figure 4.11. In short, J exhibit a higher rate of Subject-verb agreement than M, regardless of proficiency, although the trend is for an association of higher Subject-verb agreement with increasing proficiency.

Table 4.8

ANOVA of the Rate of Lack of Subject-Verb Agreement (AGR 3) by the Independent Variables of L1 and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance
Main Effects	0.16	3	0.05	4.14	0.00 **
L1	0.09	1	0.09	6.96	0.00 **
PROF	0.07	2	0.04	2.74	0.07
2-Way Interactions	0.03	2	0.01	0.99	0.37
L1 x PROF	0.03	2	0.01	0.99	0.37
Explained	0.19	5	0.04	2.88	0.02 *
Residual	2.42	204	0.01		
Total	2.83	209	0.01		

```
*    p < 0.05
**   P < 0.01
```

4.10.7 PRO-Drop and Serial Verbs

The last set of tests to be reported involve PRO-drop (PD) and serial verbs, which do not fall within the scope of the above ten hypotheses. Both measures hypothetically occur in M data more than in J data (possibly due to the greater Tp of M; see Chapter II), and should correlate to some extent with proficiency level (as in Hilles 1985, White 1984, 1985). Frequencies for each are shown in Table A; sums and means of PD for M and J, respectively, are 44 and 14, or 0.42 and 0.13. This pattern is shown in Figure 4.12 below.

Figure 4.12

**Means of PRO-Drop and Total Serial Verb Scores
Classified by Dependent Measure and Ll**
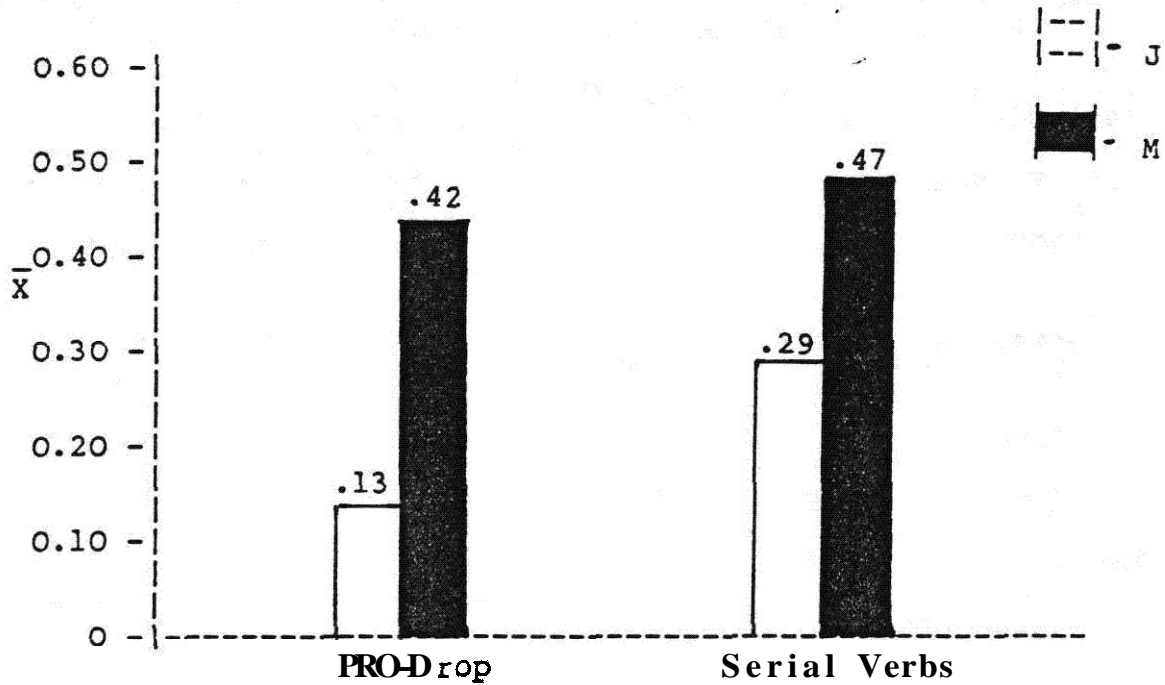
An ANOVA determined that there was a significant main effect for Ll (F=6.85, df=1/204, p < .01), as presented in Table 4.9. The effect for proficiency (means of which are plotted in Figure 4.13) was not significant, however (F=1.20, df=2/204, ns).

Table 4.9

ANOVA of the Dependent Measure of PRO_Drop by the Independent Variables of L1 and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance	
Main Effects	5.78	3	1.93	3.08	0.03	*
L1	4.29	1	4.29	6.85	0.01	**
PROF	1.50	2	0.75	1.20	0.31	
2-Way Interactions	0.60	2	0.30	0.48	0.62	
L1 x PROF	0.60	2	0.30	0.48	0.62	
Explained	6.38	5	1.28	2.04	0.07	
Residual	127.60	204	0.63			
Total	133.98	209	0.64			

*   p < 0.05
**   p < 0.01

Figure 4.13

Means of PRO_Drop and Total Serial Verb Scores Classified by Dependent Measure and Proficiency Level

113
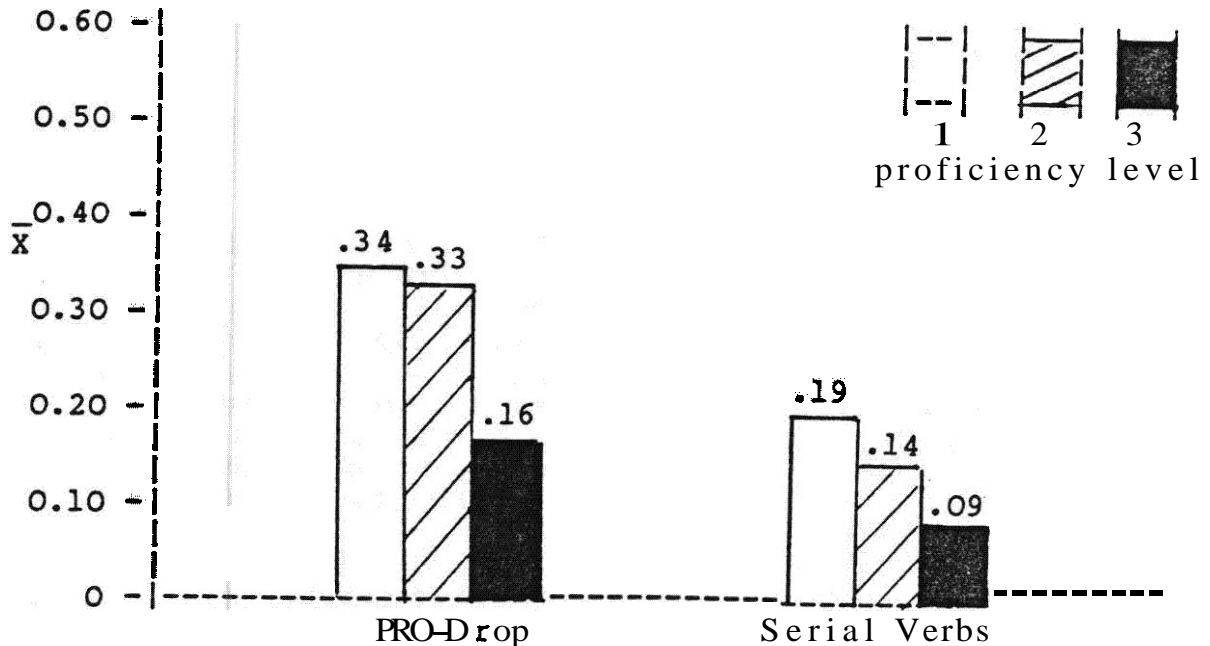
Finally, the analysis of total production of serial verbs revealed the same main effects as reported for PD: namely, there was a significant main effect for Ll (F=7.01, df=1/204, p < .01), but not for proficiency (F=1.15, df=2/204, ns); see Figures 4.12 and 4.13, and Table 4.10.

Table 4.10

ANOVA of the Dependent Measure of Total Serial Verbs by the Independent Variables of Ll and Proficiency (PROF)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	Level of Significance	
Main Effects	1.42	3	0.48	3.10	0.03	*
Ll	1.07	1	1.07	7.01	0.00	**
PROF	0.35	2	0.18	1.15	0.32	
2-Way Interactions	0.37	2	0.19	1.21.	0.30	
Ll x PROF	0.37	2	0.19	1.21	0.30	
Explained	1.80	5	0.36	2.35	0.04	*
Residual	31.20	204	0.15			
Total	33.00	209	0.16			

*   p < 0.05
** P < 0.01

4.11   Summary of Results

From the foregoing report of tests performed on the data, the analyses with significant results will be briefly summarized below (see also Table 4.11). Then, in the following chapter, the results (both significant and non-significant) will be discussed more fully, especially in terms of trends which emerged in the analyses. In most cases, these trends were consistent with the hypotheses, but they were not statistically significant; possible explanations for these results will also be offered.

The statistically significant differences across groups in the production of specific structures analyzed were as follows: (1) passives (main effects for Ll and

114

proficiency, and 2-way interaction effects **as** well, but with J production contributing the most variation to this effect: J **> M)**; (2) raw frequencies of verbs without Subject-verb agreement (main effect for **L1** only: M > **J)**; (3) ratio of verbs with Subject-verb agreement out of the total (main effect L1 only: J > **M)**; (4) ratio of verbs without Subject-verb agreement out of the total (main effect for **L1** only: M > **J)**; (5) total production of **7a-7f** (main effect for L1 only: M > **J)**; (6) PRO-drop (main effect for **L1** only: M > **J)**; and lastly, **(7)** serial verbs (main effect for **L1** only: M **> J)**.

Table 4.11

Summary of Results

Direction of Hypothesis		Statistically Supported?	Trend Supported?*
H1	**M > J** TCs 7a, 7b	No	Yes
H2	**M > J** Heavy Subjects	No	Yes
H3	M > J Existential <u>there</u>	No	No
H4	J > M Extraposition with <u>it</u>	No	Yes
H5	J > M Total Dummy Subjects	No	Yes
H6	**J > M** Sentence-initial locatives	No	(weak)
H7	**J > M** Passives	Yes	Yes
H8	**J > M** Preposed <u>if</u> clauses	No	Yes
H9	J > M Preposed <u>when</u> clauses	No	(weak)

*criterion for support of "trend" is at least 10% difference in production means

116

Table 4.11
Summary of Results (Cont'd)

Direction of Hypothesis	Statistically Supported?	Trend Supported?*
H10 Syntacticization 2a > 2b > 2c 1 > 2 > 3	No	No
If/when 1 > 2 > 3	No	Yes
Extraposition/expletives 3 > (2) > 1	No	Yes
Passives 3 > (2) > 1	Yes	Yes
7a> 7b> 7c> 7d> 7e> 7f	No	Yes
AGR 3 > AGR 1 level 1 > 3	Yes	Yes
PD/ Serial Verbs M > J; 1 > 2 > 3	Yes (M>J)	Yes

*criterion for support of "trend" is at least 10%
difference in production means

CHAPTER V

DISCUSSION

5.1 Introduction

This chapter is concerned with discussing the results presented in the preceding chapter with the objective of describing and explaining main findings in terms of the specific hypotheses which were tested. Chapter VI, the Conclusion, in turn relates the findings in the literature and in the present study to the general research questions posed in Chapter I and in Chapter III.

In Chapter IV it was reported that, by and large, the ten hypotheses did not yield statistically significant results, but nonetheless, indicated positive (albeit weak) trends in the predicted direction (see Table 4.11). The meaningfulness of the individual trends is not clear, however, since in some cases, the group differences under investigation involve low frequencies of tokens. Furthermore, the criterion for what might constitute a trend was arbitrarily set at a 10% difference in mean scores. In many cases the absolute percentage difference between compared scores was greater than 10% (see Figure 4.1 and Table A), and perhaps a larger criterion difference (say, 50% or more) would be a more valid indicator of the so-called trends.

In the following sections, unless it is explicitly stated that the trends were statistically significant, it can be assumed that the basis for quantitative judgments such as "M tended to produce more of x than J" is simply this 10% minimum difference in mean scores.

what is perhaps more interesting than the fact that there were some trends supporting the hypothesized differences between groups, is that, contrary to expectation, the individual hypotheses were generally not statistically supported. Although it can be disappointing and puzzling when hypotheses must be rejected, it also stimulates further consideration of the phenomena in question, which is a valuable part of research.

One of the most plausible explanations for the similarities across groups in terms of production of structures is that the data base is too small in some cases to test with inferential statistics, and that a more

powerful and generalizable analysis would need to include many times the number of subjects used in this study; or alternately, it would need to involve a much larger corpus of data for each subject. (And yet, it is also possible that even if significance resulted, there might not be major differences from the present findings.)

In addition to an explanation for the-results based on production frequency, there are other possible explanations to be considered. Below, therefore, the results of the individual hypotheses will again be examined and inter-preted.

Appendix C contains excerpts from J and M texts at the three levels of proficiency, as indicated. For selected subjects (whose production was interesting), sentences containing any of the structures coded for (or combinations thereof) are listed. The purpose of this is to show the range and association of the production of' some measures relative to others. It is obvious that there is a great deal of variation with regard to the (hypothesized developmental) level of some structures produced relative to the level of other structures for the same subject, such as the co-occurrence of 7a's and 7fs, or, say, the co-occurrence of passives, extraposition, and lack of Subject-verb agreement. In the discussion that follows, examples from Appendix C will be cited in several places.

5.2   Production of TCs (7a,7b)

The results for Hypotheses (1) and (2) were not signi-ficant, but it was found that M tended to produce more TCs than J (actually double for 7a and 7bb), and among them, more "heavy Subjects" as well, although it was noteworthy that both groups collectively produced quite a few TCs. Furthermore, tests for Hypothesis (10) revealed that more 7a constructions tended to be produced at levels 1 and 2 than by level 3 learners, as predicted, but that 7b (and 7bb) TCs were produced equally often by low and high level learners. The latter finding had two possible explana-tions.

The first explanation is that TCs persist in the production of subjects whose L1 are Tp (or TSp), even at high levels, and any major differences in production between M and J might be more visible at even lower profi-ciency levels (and thus lower rungs on the developmental ladder). In other words, according to Andersen's (1983) work, it is conceivable that learners whose L1 has a

structure which is similar to one in IL or in L2 (e.g., Japanese NP-NP-VP, as in 7b) will produce the L2 structure earlier than learners in whose L1 the corresponding structure is less productive, or *who* would tend instead to use a developmentally earlier structure (e.g., Mandarin NP (Topic)-VP, as in 7a). From this point of view, one might expect M at lower levels of proficiency to produce more 7a's than J at the same level, or perhaps that M would move more slowly from 7a through 7b toward 7f than J.

The second explanation is that even though the production of TCs is quantitatively almost the same between levels 1 and 3 (see Figure 4.6), there are qualitative differences between the kinds of topicalizing strategies used. That is, while heavy Subjects, left dislocation of NPs, and double Subjects occur at lower levels, by higher levels the learners are able to topicalize according to English Topic-marking conventions,

The first explanation cannot be tested with the present data, but there is supporting evidence for it in the locative developmental data to be discussed as part of Hypothesis (6). That is, the 2a frequencies, like those of 7a, are low enough to suggest that the learners in this study are all well beyond that stage of acquisition. Therefore, there is a "floor effect" of sorts in terms of production of the structures thought to be *most* developmentally basic (compared to production at more advanced stages, where there is a kind of "ceiling effect").

Indeed, even of those structures reportedly produced at the level of 7a, 2a, or even 7b, some of the so-called TCs may have been due to faulty punctuation, as *the* following sentences indicate:

1. Most high schools, in Japan are made for girls and *boys* separately (S36-J).

2. Having an education, can train a person to understand things better and *to* learn how to solve problems (S175-M).

3. The 71-year old president, Ronald Reagan wants the United States to be *the* good and old America (S20-J).

The constructions were considered in context, however, and it was assumed that comas reflected psychologically real pause markers, such as typically separate Topics from

Comments (of course, oral discourse would help to clarify this problem).

The second explanation is more easily tested, and in fact, a comparison of a subset of TCs for J and M at level 1 compared to level 3 (see Appendix C) confirmed this post hoc "hunch[n]". It appears that at level 1, most TCs involve simple, preposed (topicalized/left dislocated) NPs, and sometimes sentences or prepositional phrases with for or from. Skimming the examples shown in Appendix C, level 1 TCs seem to be less native-like than those at level 3, which include more phrases like "As to NP," "As for NP," or prepositional phrases. Compare the examples below (taken from Appendix C); of relevance is the difference between level 1 and level 3 Topic-marking, and not cross-linguistic differences.

Level 1

J: The hardest period of my life, it's I think, the time I graduated college (S54).

M: Especially the commercial, women buy anything that the TV says very good (S131).

Level 3

J: As for Kyoto and Nara, there are alot of books and advertisements on them (S37).

M: As far as a family is concerned, men and women should share everything (S175).

Of course, this is a generalization about TC differences across proficiency levels, and there are counter-examples as well; nonetheless, it does serve to explain, if tentatively, the pattern observed in TC production across proficiency levels 1 and 3.

5.3   Production of TCs with PP-NP-VP (7bb)

In connection with TC production and the move toward syntacticization, the data suggest that 7bb is an insightful category insofar as it reveals ways in which learners attempt to assign case roles (prepositions) to NPs which, presumably, at an earlier developmental stage were simple Topic NPs without (IL) case assignments. Thus 7bb might be an important and productive intermediate stage in the process of integrating Topics into English semantic and

121

syntactic sentence structure. Observe, for example, the sentences below. The structures in (b) were produced in the study (see Appendix C); (a) sentences represent a hypothetically earlier stage of production than (b); and (c) structures represent a hypothetically later stage:

1. (a) His <u>opinion</u>, he thinks as long as he can make enough money to support our family is enough.

   (b) <u>And</u> <u>for</u> <u>his</u> <u>opinion</u>, he thinks as long as he can make enough money to support our family is enough (S180-M).

   (c) <u>And</u> <u>as</u> <u>for</u> <u>his</u> <u>opinion</u>, he thinks as long as he can make enough money to support our family, it's/that's enough.

   OR <u>And</u> <u>his</u> <u>opinion</u> <u>is</u>/it is his opinion that as long as he can make enough money to support our family, it's/that's enough.

2. (a) <u>School</u>, I have seen that the student of U.S. are have so much freedom.

   (b) <u>For school</u>, I have seen that the student of U.S. are have so much freedom (S186-M).

   (c) <u>As</u> <u>for</u> <u>school</u>/<u>As</u> <u>far</u> <u>as</u> <u>school</u> <u>is</u> <u>concerned</u>, I have seen that American students have a lot of freedom.

In short, this IL use of sentence–initial preposition–al phrases accounts for the relatively high means shown in Figure 4.6 for that structure (just as it was suggested by Rutherford 1983 for Japanese preposed locatives with the preposition <u>in</u>). Moreover, its prolonged presence at higher levels of proficiencies reflects, in part, the difficulty non–native speakers have choosing suitable case frames (prepositions) for NPs, which interacts with their use of the functional strategy of topicalization.

A further difficulty for learners is that English Subjects can bear a wide range of semantic functions (compared to many other languages), such that prepositional phrases are not necessary. Consider, for example, the following pairs of sentences. The (a) sentences were produced in our data by ESL learners; the (b) sentences represent their acceptable English counterparts, which have more integrated semantic roles borne by Subjects:

122

1. (a) For example, if our human beings use TV and radio, we can easy understand that what happen today all of the countries (S33-J).

   (b) TV and radio help human beings to understand what happens everyday around the world.

2. (a) If we use car, we can go anyplace in the country (S33-J).

   (b) Cars can take us anywhere in the country. (or, alternately, we can go anywhere by car.)

3. (a) If we use washing machine, we can cut off the time instead of hand washing (S33-J).

   (b) Washing machines can cut the amount of time we spend washing.

4. (a) To Osaka a visitor will have nearly the same expression as he has on Tokyo (S37-J).

   (b) Osaka will give visitors (tourists) nearly the same impression as Tokyo.

5.4 Production of Preposed If Clauses

The immediately preceding examples included cases of if clause preposing, which leads us to the discussion of the findings for Hypotheses (8) and (9). The predicted functional use of if clauses, claimed to be more productive in J-ESL than M-ESL, and used more at lower proficiency levels than advanced ones, was generally supported: the same trend for when clauses was not as strong, however. As was shown above, the preposed if clauses were used strategically by learners to both lessen the syntactic density of sentences, and furthermore, to lessen the semantic markedness of Subjects. Other examples of the functional use of if clauses follow:

1. But if we think about the language, a language is one of the culture (S105-J).

2. If the small country was governed by the large country, it seems like to be a colonized country (S78-J).

123

3. If you **couldn't** provide a good living condition
   or the good education **that's** really cruel to
   the children **(S152-M)**.

4. If I want to find this answer, I should know
   the duties of a wife **(S116-M)**.

5. If you like two children, **it's ok**. If **you** own
   only one baby, **it's** the best **(S118-M)**.

6. If a friend cheated you, then he is not a good
   friend **(S124-M)**.

The foregoing examples illustrate the role of the
initial **if** clauses as frames for the following clauses, and
the looseness of the syntax that is associated with
constructions of this kind. The pragmatic scene-setting
function appeared to be shared by J and M learners
(although there was a tendency for J to produce more than
M), especially at the lowest levels of proficiency;
however, cross-proficiency differences paralleling those
described above for **TCs** were not apparent. This finding
for **if clause** preposing in M data does not disconfirm
findings in **Duff's** (1984) data, but it indicates that it
might be less uniquely performed by J than hypothesized in
the present study. In fact, an analysis of the use of **if**
or **when** constructions in J and M reveals that in both
languages the **if/when** clause precedes the main clause, and
there are numerous structures available to express the
conditional or temporal relationship (**Li** and Thompson 1981,
Kuno 1973). Thus, the fact that preposed **if/when** clauses
occur in J- and M-ESL should not be surprising, but it is
interesting that they are **"overproduced"** to fulfill other
functions **as** well.

Some characteristic functions of preposed **if** clauses
were that they appeared to serve as a substitute for
relative clauses or extraposition, which would entail on
the one hand more syntactically condensed Subjects, and on
the other hand, dummy Subjects with more syntactically
condensed predicates. The following sentences illustrate
this usage of **if** clauses:

1. But if you talk to a foreigner, we trust him
   when he show his nationality **(S105-J)**.
   (e.g., 'Foreigners who reveal their nationality
   can be trusted[t])

2. If **I** can understand them, **it will** be improvement for me **(S3-J).**
   (e.g., 'It would be an improvement for me to be able to understand them')

3. If you want to find **it** on the **map,** then you have to spend a long time **(S135-M).**
   (e.g., 'It takes a long time to find **it** on the map'**)**

4. If one have some knowledge in language, one will be able to communicate easily and his business will be successful **(S64-J).**
   (e.g., 'People who know more than one language will be successful . . .')

5.5 Production of **Dummy** Subjects

We will turn next to a discussion of the results for Hypotheses **(3), (4),** and **(5),** all of which deal with the production of dummy Subjects at some level. These hypotheses were included in the study primarily to address **Rutherford's** (1983). prediction that **J,** as a **language** with more grammatical word order than M due to its rigid **verb-**final status, would cause learners to produce more dummy **Subjects.(as** grammatical place holders) than M. Conversely, M, due to its relative strength as a Tp language? would not produce dummy Subjects (especially **it)** to the same extent, and would, rather, produce more marked constructions, such as heavy Subjects.

Hypothesis **(3)** anticipated a greater total production of the expletive **there** for **M,** however, on account of its claimed discourse-syntactic and developmental function in M data. This hypothesis was based on contradictory previous findings (namely, Schachter and Rutherford 1979, and Rutherford **1983),** therefore **it** is not surprising that the results pointed to a roughly equivalent quantity of production across the two language groups for this structure. With respect to the possible role of **there** as a developmental trigger that was mentioned in earlier chapters, **it** was clear that the structure was already well integrated into the grammars **of** level 1 learners, that **it** peaked at level 2, and declined in frequency somewhat by level 3. This particular finding mirrors other sets of results **(e.g.,** production **of 7a TCs** and 2a **LOCs,** compared to 7b, and 2b, respectively? and the overall production **of** verbs without Subject-verb agreement compared to those marked for agreement). In other words? if **there** or

125

agreement are indeed triggers in acquisition, then the present data represents a level which exceeds any stage at which this shift would occur.

Hypothesis (4), by comparison, revealed the predicted tendency for J to produce more extraposition with _it_ than M, although the tests were not statistically significant. Extraposition occurred at all levels, like there, and followed the same general pattern of peaking at level 2 and declining thereafter. The kind of IL extraposition usage ("overproduction[n]") that has been reported in the literature was evidenced by both J and M, as the following examples illustrate, although extraposition of this type was, on the whole, relatively infrequent.

1. Secondarily **it** [non-referential] is a good way to make many plantations of vegetables **(S9-J, level 3)**.

2. **It is** welcome that Japanese people have increasing opportunities to visit another Asian countries **(S17-J, level 3)**

3. **It** [non-referential] have been become very convenient our daily lives **(S33-J, level 1)**.

4. **It** has large Japanese garden behind of the shrine **(S68-J, level 1)**.

5. **It is** hoped to develop the unfamiliar kinds of fishes **(S67-J, level 3)**.

6a. **It** can be said **"human,"** when people have these three independence.

b. **It** cannot be said **"men** should be given the first choice for college entrance, for they are the chief wage-earners[n] **(S72-J, level 2)**.

7. Since then **it** have passed ten years **(S86-J, level 1)**.

8. **It is** the first reason **I** want to apply an American school, to learn more **(SS113-M, level 2)**.

9. **It** comes another problem to me to overcome, **it** is the problem of normal courses **(S138-M, level 1)**.

10. In other side, it has cause a very serious social problem in Taiwan is the student cannot get high score in the competition (S173-M, level 2).

Hypothesis (5) dealt with the combined effect of the dummy Subjects it (in extraposed constructions) and there, and it was reported that J produced about 10% more of the pooled structures than M, although the barchart in Figure 4.7 shows that both Ll groups were productive in this respect. This finding is perhaps not strong enough to discredit Rutherford's notion that J are more sensitive to filling Subject position than M, and observing the 'Heavy-to-the-Right" principle (Mallinson and Blake 1981). Rather, what it suggests is that there might in fact be a transfer-of-training effect based on the nature of the writing task, which reinforces the quantity of dummy Subjects generated by both J and M learners. This suspicion was already mentioned in describing the results of the Duff (1984) study, in which conversational data by M and J learners evidenced little if any dummy Subject production. On the other hand, a counter-argument is that Duff's (1984) subjects were all at level 1, whereas, as it was noted above, dummy Subject production appears to peak at level 2.

There is, nevertheless, good reason to believe that learners are aware of the need to use impersonal agents in formal written English (especially in technical writing) as opposed to personal pronouns. It is, then, slightly peculiar that the same trend (for comparable production for J and M) would not be evidenced in overall production of passives, which are also associated with the style and rhetoric of written discourse.

A more controversial explanation, perhaps, is that the theoretical basis for the prediction that J produce more dummy Subjects than M (cf. Rutherford 1983) is ill-conceived, and that the partial GWO status of J (but not M) has no bearing on dummy Subject production. [note 6] In fact, if J were to produce significantly more dummy Subjects than M, two other reasons could be suggested (i.e., in addition to the the discourse-functional role that Schachter and Rutherford 1979 documented for e.g., it): (1) Mirror image rules operate to transform J sentences of the type "[Sentence] no da" to ESL "It is (the case) that [sentence]," or J " [Sentence] to iwareru" to ESL "It is said that [sentence]," and so forth (Smith 1978, Kuno 1973); and this effect is particularly visible in expository writing with topics of a generic nature. (2) It

127

is the Sp status of J (as opposed to GWO) which brings about the tendency for dummy Subject production, as in extraposition with it. Thus, the flexibility of the language in terms of word order scrambling has less bearing on IL dummy Subject production than the basicness of the Subject function in the L1.

5.6 production of Passives

In addressing the findings of Hypothesis (7), it was reported that J produced significantly more passives than M, and that there was an interaction effect for proficiency. Further analysis revealed that the proficiency effect was most pronounced and significant for J, but not M. That is to say that M not only produced many fewer passives than J, but that across proficiency levels, M data reflected the pattern 2 > 3 > 1, the same pattern as was reported for extraposition and there production (see Figure 4.5). J data, by contrast, displayed a very strong directional progression: 1 < 2 < 3, with a particularly noticeable leap in production between levels 2 and 3. Several possible explanations might be considered for this major difference.

First, although both J and M are said to have passives with affective or, adversive semantic content, passives in J (1) are more productive than in M; (2) are more syntacticized than they are in M; (3) involve a wider range of semantic features than those in M; (4) typically occur in generic, neutral statements of fact; and (5) are an important reflection of the basicness of Subjects (i.e., Sp).

It is beyond the scope of this discussion to present a complete description of J and M passives (see e.g., Howard and Niyekawa-Howard 1976, and Li and Thompson 1981). However, related to point (2) above, it appears that J passives involve the verbal suffix -rare plus the case marker ni, as in

Kodomo wa okaasan ni shika-rare-ta.
child TOP mother by scold-PASS-ed
'The child was scolded by his mother[8]

Mandarin, on the other hand, forms passives by creating a rather loose string of constituents which includes the passive marker bei: i.e., NP1 bei NP2 verb, as in:

```
Ta     bei  jiejie         ma    le
3 sg PASS elder sister scold ASP
'S/he was scolded by (his/her) sister'
                       (Li and Thompson 1981:492)
```

Concerning point (3) above, whereas J has both affective and neutral passives (as well as indirect and direct passives), as in (a) and (b) below (Kuno 1973:302), M passives are much more restricted in semantic scope.

```
a.   Kono ie    wa 1960-nen ni tate-rare-ta
     this house TOP    year in build-PASS-ed
     'This house was built in 1960' (neutral)

b.   John wa niwa no sugu  mae  ni
               yard 's right front in

     ie   o    tate-rare-ta
     house ACC build-PASS-ed

     'John had a house built on him right in  front
     of (his) yard' (adversive)
```

Another explanation is that J tend also to produce more cases of extraposition, and since passives figure prominently in J-ESL extraposition (i.e., "It is believed that," "It is hoped that," "It is expected that[n]"), there is. an even greater main effect for L1. Statements of fact in J are often made with passives of the type in (a) above, with an impersonal (often sentential) surface Subject, and this usage might be expected to transfer in L2 expository writing.

A third explanation is, as Rutherford claims, that the combined factors of GWO/PWO and TSp in J typology serve to promote the use of passives, relative to M.

Fourth, it is possible that there is training effect for the production of passives in Japanese EFL (English as a foreign language) education, more so than in the Chinese EFL context, but we have no evidence for this supposition.

A final note concerning passives is that there did not appear to be interesting qualitative differences in the production across groups, but rather the primary distinction to be made was quantitative.

5.7   Production of Locatives (2a, 2b, 2c)

It was mentioned above that the locative data that were the focus of Hypothesis (6) seemed to exhibit a "floor effect" at the lowest stage (2a), suggesting that the subjects were already well beyond that point of acquisition. Here, we will consider further the outcomes in terms of Rutherford's model, and the role and development of locatives and existentials in J data according to his claims.

The findings for locative production, summarized in Table 4.11, showed little support for the hypothesis based on Rutherford's developmental prediction (e.g., Table 2.2), either statistically or in terms of general trends. Although it is really not possible to speak to differences at the-stage of 2a, due to low numbers, it is clear that J/M means for 2b (0.41/0.37) are not nearly as distinct as one would have expected, and neither is the combined total of locatives and temporals (2b+2bb) very different across Lls. Related to this, the structure in 2c which was proposed to evolve out of the locative in 2b, and thus occur at higher proficiency levels as opposed to lower ones, in fact followed the opposite order in terms of production frequencies: 1 > 2 > 3.

It is not completely clear how these findings are to be interpreted. On the one hand, there are many obvious, typical instances of the locative phrase "In Japan, NP/there . . . " in the J data (see Appendix C); on the other hand, there appear to be nearly as many locatives in the M data, but they are slightly less obvious. This might be due to the diversity of locative expressions produced in M; that is, whereas J tended to use locative expressions with specific place names, M produced more generic, abstract locatives, such as "In society,' "In civilization," and particularly the phrase "In a/the family." Therefore, the use of preposed locatives is initially less apparent for M, and there may be subtle ethnolinguistic factors underlying the choices (i.e., J-ESL "In Japan" vs. M-ESL "In a family").

An explanation for the rather high frequency of (2b) locative production by either J or M is that the task and the circumstances of the task lend themselves to production of this form. In other words, since the subjects were asked to write about general topics concerning cross-cultural differences in customs and values, it is not surprising that the topicalization of locatives should take place. Recall also that Topics often bear a contrastive

semantic feature, and that in comparing American and Japanese or Chinese ways, for example, this construction is an effective rhetorical choice.

In terms of circumstances, many subjects writing the compositions are recent arrivals from their respective countries, experiencing culture shock and various first impressions. For this reason, they are inclined to compare and contrast these different locations.

Again, to speak more authoritatively about the true differences between M and J use of locatives, it would be useful to examine a large corpus of oral production data (as in Duff 1984) including narratives as well as expository discourse (e.g., picture descriptions) from comparable groups of learners. Moreover, it would be important to collect data from subjects at a much lower level, in order to address the question of the production of 2a with respect to 2b in terms of a developmental sequence. Only then, in fact, would we also be able to test differences between J and M in terms of the initial point of development on this cline for each, and concerning the pace at which the stages are traversed.

Lastly, concerning locatives, Kuno (1973) suggests that J, M, and English existentials are structurally quite similar, with the basic word order being a preposed (definite/generic) locative, followed by a sentence with an indefinite NP and an existential verb. In deriving existentials of the type corresponding to our 2b, however, he posits first a construction such as 2c ("There are . . . locative"), which in its derivation is followed by the 2b construction with the pragmatically preposed adverbial. It is not possible to expand on Kuno's (1973) argumentation here, but it is worth exploring in future studies whether Rutherford's sequence of 2a > 2b > 2c should not perhaps be 2a > 2c > 2b, instead. The latter series would not only be consistent with the derivation (and possible developmental complexity) of existentials in Kuno's work, but also would account for the pattern of locative production observed in the present study for both J and M.

5.8 Production Reflecting Syntacticization of Topic

Hypothesis (10) examined the effect of proficiency on the syntacticization of Topic. Some of the issues related to proficiency for certain structures have been addressed in the preceding text; what is of greatest concern here is

a fuller explanation for the pattern observed in Figure 4.6.

According to the hypothetical pattern sketched in Figure 4.7, the results comprise a very "top heavy" set of data; clearly, if Figure 4.6 were a seesaw, the right side would outweigh the left side considerably. In other words, the process of syntacticization of Topic is well underway for the learners tested, and in some cases it is complete, although there tends to be some "free variation" of structures at times (see Appendix C). Incorporating the measures of agreement (e.g., AGR 1 and AGR 3) is very important and promising as another index of syntacticization in SLA; eventually the scale conceptualized in Figure 4.6 will need to be expanded, though, or drafted along different intersecting or possibly parallel axes with other manifestations of syntacticization, such as the development of relative clause formation and article usage. As the scale now stands, it may give too much weight to the (proposed) mid-way stages, especially 7c and 7d; on the other hand, if there had been more structures of that kind produced, their actual developmental role night have been clearer.

The production of lack of Subject-verb agreement (AGR 3) revealed significant differences between J and M; in short, M produced more verbs without agreement than J, although the M data (see Figure 4.9) exhibit a striking drop in mean scores on this measure between levels 2 and 3, with level 3 means even lower than those for J at the same level. Is it possible that between levels 2 and 3 this shift in agreement is characteristic for M but that J, whose L1 presumably has a more developed (syntacticized) VP, adopt English Subject-verb agreement much more quickly than M? This point deserves further study in a theoretical framework that can explain the development, as opposed to just documenting morpheme acquisition as earlier studies in this field did. Furthermore, in studying the development of the ESL VP in J and M learners, it would be useful to look at the development of tense and aspect as well.

5.9   Production of PRO-Drop

Finally, PRO-drop and serial verbs are two structures which yielded significant differences across L1 groups, as predicted. In what follows, we will discuss qualitiative and quantitative differences in the production of PRO-drop, in particular, which is thought to be a parameter in UG and a typological parameter. Serial verbs, on the contrary,

132

are arguably a more language-specific syntactic phenomenon which represents in the IL data a fairly transparent transfer effect from the M Ll.

The results for PRO-drop were that M evidenced approximately three times as much PD as J. For the entire group, the progression in terms of -PD relative to +PD was, roughly, level 1 = 2 > 3. Qualitative differences across Ll groups suggested that in the M data there was more ellipsis of third-person (3p) generic, impersonal, or non-referential Subject pronouns, whereas in J, there tended to be more first-person (1p) Subject ellipsis. While the kinds of PD that occur can, in the case of M, be largely attributed to either the non-referential nature of the pronouns, or their pragmatic recoverability (or even their relatively high frequency in conversational English), it appears that in the case of J, PD occurred more in complex syntactic constructions than in simple sentences. This indicates that J might be attempting to apply (here, "over-apply") higher level English rules of, for example, Subject PD in coordinate constructions.

The finding that 3p referents are easily deleted is not a new one; it is supported by the work of Chen (1984) and Li and Thompson (1979) for M as Ll, and by Zobl (1984) for ESL. These linguists claim that 3p PD, and especially non-referential or generic PD, is productive (here, in M-ESL).

In addition to the quantitative difference for M vis-à-vis J, the qualitative difference hints in another way at the sensitivity of J to filling the Subject position in ESL, even though ellipsis is common in the Ll (Hinds 1978, 1983).

Topic chains were neither common nor particularly noteworthy in our data (contrary to expectation), with the exception of one M learner (S185), who used a listing strategy akin to that described in Schumann's (forthcoming) work, which was outlined in Chapter II (Section 2.3.7.1). Another M learner, (S109), used a strategy of introducing sentence Topics, then immediately deleting them, and this is found in Appendix C (M, Level 1, S109), although the remainder of the text in which this pattern is displayed did not occur within the 12 T-units analyzed, hence it is not in the present analysis.

Below, examples of M and J PD are cited, and the listing strategy referred to above is presented in Appendix C (M, level 2, S185). Notice that a couple of cases of

Subject-verb inversion occur in the examples (3, 12), and these are thought to be one of several composite features subsumed by the PD parameter. Another feature is that of Chinese-style "surface" or pseudo-passives, in which there is a Topic and predicate, but an ellipted agent (6, 13).

Mandarin PD

1. ( ) True, that television has negative effect (S165).

2. But basically ( ) depends on the husband and the wife (S182).

3. In the summer (INV) was the happiest time (S193).

4. The third, ( ) should keep the house clean as possible all the time (S182).

5. They don't have the kind of sex of baby that they wish for and ( ) keeps on getting preganut (S204).

6. More than this limitation ( ) will have to pay a certain tax.

7. ( ) Hope I could improve the relationship more so we could discuss more freely (S179).

8. ( ) Can thus broaden my scope (S117).

Japanese PD

9. I went to the college only by one reason that to ( ) wanted to become such two men (S1).

10. Before you visit my country, I would suggest that ( ) read and learn about my country Japan (S12).

11. Especially in my study field ( ) is minor part (S2).

12. In Hawaii always blow wind named "Traffic Wind" (S12).

13. Thesis, words Or something like that ( ) had to translate into Japanese (S12).

134

5.10   Summary of Chapter **V**

The   aim of this chapter was to examine   more   closely
the   trends or significant differences reported in   Chapter
IV.   In discussing the reported differences, several points
were   raised concerning the interpretation of the findings.
These are summarized below:

(1)   The   frequency of production   of   structures   was
relatively **low,** and   this   affected the **degree** to   which
generalizable differences could be claimed.

(2)   The relatively high proficiency level of   learners
whose   lowest   level was **TOEFL** 450 created a kind of   floor
effect at that end of the scale,   such that the   structures
predicted   to   occur at the lowest proficiency levels   (and
differentially across groups in some cases) did not occur.

(3)   The expository nature of the   composition   topics
promotes   certain   kinds   of   structures   **(e.g., preposed
locatives,** and   generic statements with extraposition   and
passives)   which   might   not occur to the   same   extent   in
personal   narratives   or other kinds of written   or   spoken
discourse.

(4)   The   writing task encourages the   production   of
dummy Subjects   and   passives due to a   possible   training
effect,   and these constructions would likely occur less in
oral production.

(5)   In   addition   to   the   reported   quantitative
differences, qualitative differences between   or   across
groups were observed and highlighted.

CEAPTER VI

CONCLUSION

6.1   Introduction

The research questions presented in Chapter I and Chapter III for the thesis in general and for the study were as follows:

(1)   What is the effect of the independent variables of L1 (J or M) and proficiency level on the dependent variable of syntacticization of Topic?

(2)   Secondly, to what degree are these effects accurately captured by Rutherford's (1983) model?

(3)   Eow is the claimed transfer effect of Tp consistent with other constraints on language development?

The purpose of this chapter is (1) to summarize the answers to these questions; (2) to state the limitations of the present research; (3) to consider directions which future research in this area might take; and finally? (4) to suggest possible pedagogical implications that can be derived from this work.

6.2   Answers to the Research Questions

The first research question concerns the differences that were observed in J and M data which could be accounted for, in part, by typological differences (assuming that the typological distinctions of Li and Thompson 1976 and Thompson 1978 are valid).   In addition, it addresses the interaction of L1 with proficiency level.

. Generally speaking? the findings revealed a (non-significant) tendency for J and M data to take different shapes? such as: (1) M evidencing more TCs at the 7a level, and of those, more heavy Subjects than J; (2) J producing more extraposition with it than M, and possibly related to this, also producing more passives than M as well; (3) M performing more PD than J, and furthermore, generating more verbs without Subject–verb agreement than J; and (4) J, on

the other hand, using about 20% more preposed if clauses than M.

Those features reported above appear to be related to a large extent to the typological differences that exist between a so-called pure Tp language with PWO (i.e., M), and a less pure Tp language (TSp) with both PWO and GWO (i.e., J). However, these tendencies were generally not statistically significant, and thus, caution must be exercised in interpreting and generalizing them to other populations.

Of course, the theoretical underpinnings of these typological assumptions could also be challenged, and one could speculate that the lack of significant cross-linguistic (cross-typological) differences observed for a number of the features analyzed might also suggest that the typological distinction between borderline cases, such as J (TSp, GWO/PWO), with more strongly typed languages, such as M (Tp, PWO), is marginal and empirically unverifiable. There was also some question as to the true effect of the PWO/GWO paradigm on the production of dummy Subjects in IL.

Regarding the role of proficiency, there appeared in some cases to be an interaction with L1, but for the most part, it was L1 that was the more robust independent variable in terms of a statistically significant main effect. To the extent that the TOEFL test yields a reliable standardization of scores, and insofar as it represents a valid measure of global proficiency that can be correlated with an index of syntacticization, this finding suggests that the subjects in the present study had already reached a level of proficiency beyond which there was "free variation" among many of the structures produced.

Turning to the second research question, it was reported that while Rutherford's (1983) model captured some of the syntacticization effects that were hypothesized, and especially so for the sequence proposed in Table 2.1, it was less accurate in conceptualizing the development of locatives for J (as in Table 2.2). What can be gleaned from this is that (1) his model needs to be retested with a wider range of learners (in terms of both L1 and proficiency level) and using a wider range of modalities (e.g., spoken, comprehension) as well: and/or (2) his model needs to be modified on the basis of quantitative research findings. Despite the limited scope of the present study, and the ambiguity of some of the results, it was recommended that Rutherford's model include a measure of Subject-verb agreement, expand the treatment of relative

clause formation, and possibly, revise the developmental sequence posited for J production of locatives to 2a > 2c > 2b. The development of existentials in locative constructions might, in fact, be best incorporated in the larger model as a different but related aspect of existential development for both J and M.

The third research question was addressed in Chapters I and 11, where it was reported that features associated with Tp and with the so-called "pragmatic" or "pre-syntactic" mode of communication (Givon 1984) were observed across a wide range of developmental data. Moreover, the discussion in Part I of Chapter II predicted and accounted for the persistence of some kinds of Tp constructions (e.g., TCs) even at higher levels of proficiency, due to their compatibility with other constraints and factors in language acquisition and language production, such as the perception of typological distance and of whether a given (L1) structure is relatively unmarked, and thus a "reasonable" or "logical" entity in IL. It was noted that the frequency of structures in L2 or IL input, and the compatibility of IL forms with these structures were other factors affecting the degree to which certain forms could be predicted to occur in IL.

This background understanding of constraints on transfer and natural IL development can account for some of the structures observed in both J and M data: e.g., TCs, preposed locatives, and if clauses with a topicalizing function. Other constructions and operations that were observed, such as PD, existentials with there, and extraposition with it were considered more indirect effects of the general sensitivity of the L1s in question to the functions and basicness of Subject. Again, though, many of the predictions and explanations for transfer effects advanced in Chapter II must be tested with a wider range of data in the future.

6.3 Limitations of the Study

Several points will be listed below as limitations of the study conducted in this thesis. Most of these result from constraints imposed by "convenience sampling," and constraints on time and other resources, but nonetheless, they must be recognized as potential threats to the generalizability of the findings.

1. The small number of subjects, and small amount of data available for analysis in each composition.

2. Differences in subjects' experience in English, based on length of previous instruction (primarily EFL), and length of residence in United States.

3. The limited range in subjects* proficiency between TOEFL scores 450–599; thus the relative paucity of structures produced at the lowest levels of the syntacticization indices used here.

4. Possible differences in conditions under which compositions were written, since the data were produced between 1980–1985.

5. The effect of composition topics on structures produced.

6. The differences in production based on written versus spoken discourse, while not a problem in replicating Rutherford's work, poses problems in generalizing to overall IL development.

7. The initial problems in inter–rater reliability testing related to the identification of potentially codeable structures (particularly true with passives and extraposition); and the amount of judgment or inferencing required in some categories (e.g., 7a, 7b).

6.4   Directions for Future Research

In spite of the limitations of this study, it paves the way for many kinds of future research. First of all, it would be worthwhile to compare the reported findings with an analysis of only the first sentence in each of the compositions (as Rutherford 1983 did) to see whether there were significant differences in the kinds of structures produced (e.g., heavy Subjects) based on the respective methods.

Secondly, an interesting study would compare spontaneous oral production with written production data, based on tasks that were held constant. One would predict that there would be more TCs and other topicalizing/

139

pragmatic strategies used in the former case, and conversely, higher frequencies of production for dummy Subjects, passives, and Subject-verb agreement in the latter case (note, though$_I$ that there are some problems with determining agreement in oral production due to the possibility of confounding IL phonological variables). .

A third kind of study would examine the longitudinal development of syntacticization in J and M ESL, from very low proficiency levels (as in Huebner's 1982 study). Related to this, the variable of naturalistic versus instructed acquisition (plus or minus literacy) would be an interesting factor to examine; one could hypothesize differences across language learning type in the syntacticization processes, the speed of development, and so on. That is, with the intervention of explicit instruction, learners might acquire the coding and behavioral properties of Subjects more quickly than naturalistic acquirers. This research question could, of course, be addressed in a cross-sectional design as well, although perhaps not as satisfactorily as with a longitudinal design.

Fourth, although in the present research J and M served as control groups for each other, future work should compare J and M production with other groups of learners (and possibly with normative L1 production data as well). For example, using the Tp/Sp typological distinction$_I$ it would be useful to examine more closely cross-linguistic/typological differences across the four proposed combinations: Tp, TSp, Sp, and -TSp. One problem in dealing with research involving learners of -TSp languages (e.g., Tagalog$_I$ Ilocano) is that there is much theoretical controversy among linguists as to the-status of the notions of Subject$_I$ Topic, PRO-drop" and so on in languages of this type; at the same time, IL analyses of learners with these L1s might shed light on some of the issues.

Lastly, as more SLA research turns to investigations of IL development in languages other than English, it will be interesting to examine differences in developmental stages and strategies observed in, say, speakers of an Sp L1 faced with the task of learning a Tp L2, or speakers of a Tp L1 learning a TSp L2, and so forth. It would be necessary in such studies, as in the present endeavor$_I$ to formulate predictions based on markedness theory$_I$ psycho-typological factors$_I$ tendencies toward transparency$_I$ and so on, such that the processes were not just assumed to be bidirectional from a given L1 to ESL and from English L1 to a given L2.

6.5   Pedagogical Implications

The final discussion of this chapter concerns the need to specify the value of this area of research for L2 teachers. First, presenting descriptions, explanations, and predictions for the kinds of processes learners will use as they gradually **syntacticize** an IL grammar should help teachers to understand not only why learners produce what they do, but also how their production is part of a much larger IL system in which there are many interrelated components. Thus, PRO-drop, article deletion, and **TCs** are to be viewed as closely related features which might co-occur in IL in a highly systematized way (recall Huebner 1979). The same has been claimed of extraposition and existentials (Schachter and Rutherford 1979, and Rutherford 1983). Again, with this knowledge, and with some background as to the universal nature of many of these developmental processes, a teacher will be more aware of learners' strategies and progress.

With regard to specific steps teachers might take to speed up the syntacticization process, there is much that researchers have yet to learn and to contribute in this area. However, it can be assumed that with a focus on communication and discourse-level language use (as opposed to a focus at the level of lexicon or isolated sentences), learners will be exposed to discourse-level syntactic devices in the **L2,** and to how discourse coherence is possible through correct use of syntacticized Topics. Rutherford **(1977),** for example, outlines exercises in which learners must connect an assortment of sentences (propositions) in such a way that given and new information must be identified and reflected in the correct choice of sentences which make up a text.

While 'it is not clear what effect explicit language instruction might have with regard to the development of **Subject-verb** agreement, dummy Subject use, and so on, it is fairly obvious that by contextualizing learners' language practice, they will have more realistic opportunities to generate and interact with language at the **discourse-**syntactic level; this will also give them opportunities to produce the range of structures examined by Givon (1984) related to establishing and maintaining Topics in discourse.

There is much to be learned by applied linguists and by classroom teachers concerning the role and effects of explicit (pedagogical) grammar instruction and of different types of classroom interaction in SLA. Likewise, there are

many aspects related to the processes and constraints in IL development that need to be explored further. The research undertaken in this thesis represents an attempt to understand some of these factors in **SLA** a little better.

NOTES

1

   This text is taken from the linguistics term paper of a Japanese university student who, at the time, was unfamiliar with either Rutherford's work or the present study. It seemed clear that the phenomenon the writer was describing was linked with that under scrutiny here, although the general context of his discussion is more concerned with problems of discourse planning in spontaneous oral narrative.

   The gist of the introspection is strikingly similar to a phenomenon which Chafe (1976) has documented. Chafe (pp. 51-53), citing examples from Caddo, an Amerindian language spoken in Oklahoma, addresses the problem of processing constraints on unplanned discourse, and the resulting hesitation phenomenon (and TCs):

> I would suppose that because of this mutual interdependence [of choice of case frame for a sentence and choice of one noun included in the case frame as Subject of the sentence], a speaker is able to think simultaneously of the most effective framework of cases to express what he has in mind and the most effective way to package it in terms of subject. It is not unusual, however, for speakers to depart from this simultaneity by choosing--and in fact uttering-- the subject before the case frame has been chosen.

> ▪ ▪ ▪ One might think of calling such prematurely chosen subjects topics, or even speculating that the origin of topics as distinct from subjects lies in this kind of aberration in the timing of the processes of sentence construction. In short, a topic would be--or might have originated as--a subject which is chosen too soon and not smoothly integrated into the following sentence.

   In such an analysis, it would be much more likely to find such "hesitation" phenomena and resultant TCs in conversational discourse than written; however, Rutherford (1983), who examined written discourse, describes Japanese-ESL development in terms of the development of existentials (and syntacticized Subjects) out of preposed locatives with a topicalizing function, as in the example cited here.

2

Magretta (1977:149-151) distinguishes Topic marking as "weak", "intermediate" and "strong," as the following classification illustrates:

<u>Weak</u> <u>Topic</u> <u>Marking</u>

1. Deep structure subject
   John bought the records.

<u>Intermediate</u> <u>Topic</u> <u>Marking</u>

1. Passive (**NP** preposing)
   The cows were milked by the farmer.

2. Subject raising
   That amplifier appeared to be defective.

3. <u>Tough</u> movement
   Sonatas are easy to play on the violin.
   This violin is easy to play sonatas on.

4. Extraposition
   It was a waste **of** time to grade those papers.

5. Cleft sentence
   It was **Jack** who broke his arm.

<u>Strong</u> <u>Topic</u> <u>Marking</u>

1. **Topicalization**
   These letters we **haven't** looked at yet.

2. Left dislocation and <u>As-for</u> phrase
   (As for) that movie, I know you **won't** like **it.**

3. VP preposing
   (She wanted to pass the **exam,**) and pass it she has.

4. Participle preposing
   Sitting in the corner was the chairman-elect.

5. Comparative **substitution**
   More important has been the growth of the economy.

6. Negative constituent preposing
   Never have I seen such a crowd.

144

7. Directional adverb preposing
    Out of the house stepped the heavily disguised midget.

8. Prepositional phrase substitution
    In the box sat a tiny kitten.

9. Adverb dislocation
    On several evenings we saw John at the library.

3

Apparently, in some dialects of English these constructions are not uncommon, however.

4

Mallinson and Blake (1981:151) point out that due to processing constraints and short-term memory, the universal tendency across languages is for:

(a) more topical (as opposed to non-topical) material to appear at the beginning of the clause; i.e., the "Topic-to-the-Left" principle.

(b) heavier material (as opposed to lighter) to appear at the end of the clause; i.e., the "Heavy-to-the-Right" principle.

5

Gordon (1985) also discusses the form-to-function relationship of conditionals in a case study of a Japanese learner of ESL. She notes that the learner in question had acquired some of the conditional functions of *if* clauses before certain others, and meanwhile, used *if* constructions to accomplish other discourse functions as well.

6

I owe this observation to Professor Jeanne Gibson.

INTER-RATER RELIABILITY TESTING

1. T-units

A T-unit is defined by Hunt (1977:93) as:

> a single main clause (or independent clause, if
> you wish) plus whatever other subordinate clauses
> or non-clauses are attached to, or embedded
> within, that one main clause.

The conjunctions and, but, so start new T-units,
unless the following T-unit has a Subject which is ellipted
through coordination with the previous sentence. Any
subordinators, such as because, since, when, which
introduce adjuncts or subordinate clauses to the main
clause, are not considered part of a new T-unit unless they
are bounded by terminal puntuation from the main clause.
Even if the resulting T-unit would be considered a fragment
in Standard English, if it contains a verb and it has the
form of a sentence, it will be counted as a T-unit. In a
coordinate construction in which the second successive
Subject is not coreferential to the first, or is ellipted
when it should not be, the second clause can be considered
an independent T-unit (with Pro-drop).

As Hunt (1970:4) writes:

> the T-unit is minimal and terminable. *Any* complex
> or simple sentence would be one T-unit, but any
> compound or compound-complex sentence would
> consist of two or more T-units.

This convention is illustrated in Hunt (1970:4-5) by the
following example of coding:

> I like the movie we saw about Moby Dick the white
> whale / the captain said if you can kill the white
> whale Moby Dick I will give this gold to the one
> that can do it / and it is worth sixteen dollars /
> they tried and tried / but while they were trying
> they killed a whale and used the oil for the lamps /
> they almost caught the whale /

Another example, from ESL learners, is as follows:

(1) In China, many kind of TV program showing /
(2) and the children watching them. / **(3)** Because
can learn many things from TV./ **(4)** [Sometimes
the mothers, big. trouble with children watching
TV.] **(5)** Can't get the children to do his
homework / **(6)** and like to make their mothers
angry sometimes./

Numbers **(1)**, **(2)**, **(3)**, and (5) are considered separate
T-units above. Number **(4)**, however, is not considered a T-
unit because it does not contain a verb (only a nominal
predicate). Therefore, it is counted as a non-syntactic
unit **(NS)** and is tallied separately. Only those **NSs**
which occur within the twelve T-units analyzed per
composition will be tallied. (6) is an example of a T-unit
with null-Subject, since the logical Subject in **(6)** is
children, whereas that of **(5)** is mothers.

Coding: Mark T-units with slashes / /, and non-syntactic
units **(NSs)** with square brackets [ ].

**By convention, code complex sentences with more than one
independent clause as follows:

He said that he wanted to buy a bike / and he was
going to do so immediately. / John will go home /
and Mary will go to work if it rains./

2.a. (In) locative-VP

This refers to sentence or clause initial locatives,
with the preposition in or an ellipted preposition in. In
the case of the ellipted preposition, there must be
evidence that the locative element is not a grammatical
Subject; i.e., if there is a pause between it and the
predicate, or often the verb got instead of has, it can be
considered more the Topic than the Subject. Constructions
like "Japan has ■ ■ ." are Subject-predicates, and are not
counted here.

(1) Japan, got lot of problem nowadays.
(2) In China in the city got this kind of problem.
(3) In my country, hasn't army, navy and air force.
(4) In this case is one is say that television is good.
(5) In my family, work hard everyday.

Coding: above the nominal (in-NP), write **2a.**

147

Note: Sometimes there is the problem of deciding what whether an element is locative or not.   NPs such as home and family can be considered abstract locatives, and will thus be included in this category.   Temporals such as youth, childhood, free time, and future will be coded separately as 2aa if they occur as Topics, or with the preposition in; they are coded because they are structurally similar and have much the same pragmatic function as locatives in that they frame the predication.

2.b.   Locative with in + sentence with there or NP

In these sentences, there is an initial locative and this is followed by a main clause with a Subject and predicate.

(1)   In Japan;   wife is expected to stay at home to take care of children.

(2)   In Japan people are expected to act as others do.

(3)   In Japan, many people,   both men and women, smoke it wherever they are.

(4)   In America, there are many kinds of people.

Coding: Above the locative phrase, mark 2b.   Again, any similar types of constructions in which the initial phrase is a temporal rather than a locative, but reflects the same pattern, e.g., 'In future, we . . .," can be coded 2bb.   Do not code "In my opinion,' "In my view," or "In addition,[n] though, as either type; these are idiomatic or formulaic.

Note: there expressions may be double-coded as 2b (or 2bb) and a category in section 7, if it is appropriate.

2.c.   Existential there + VP + locative in

For this category,   there is an initial there phrase followed by a locative.   Usually there is followed by a BE verb,   but sometimes in IL usage,   it will be followed by have.   Again, any other possibilities which do not conform exactly to this structure (temporals) can be coded 2cc.

(1)   There were always many people living in the small town surrounded by mountains and had no way to communicate with others.

148

(2) There is no bird in the tree this morning.

(3) There are many kinds of fishes in the sea of the world.

(4) Because, there are many Japanese in U.S.C.

(5) There are many old famous temples and shrines in Kyoto.

<u>Coding</u>: underline the <u>there</u> construction and code it with 2c.

Note: Constructions coded 2c (or 2cc) may also be double coded in section 7.

3. Sentence–initial <u>if</u> constructions

Sentence–initial means that the <u>if</u> construction precedes the main clause; indeed, in some texts, you will find that learners consistently separate the <u>if</u> clause and the main clause as two separate sentences, and therefore, by convention, two separate T–units. However, for our purposes, the <u>if</u> clause need not be in the same T–unit as the main clause, but it must precede the main clause. Clauses with <u>even if</u> or <u>what if</u> can be included.

(1) If you can teach a student in their preference way; I think that's the best way to teach.

(2) If we don't walk this course, we are treated as a different type of persons.

(3) If people know at least one foreign language, they can make communication with foreigners.

(4) If the population of the Earth continues to increase at the present rate, it is inevitable that we will become overpopulated to support life.

<u>Coding</u>: circle the <u>if</u> and put a "3" above it.

149

4. Sentence-initial **when** constuction

The same coding convention applies to **when** as to _if_.

   (1) But when it comes to the daily home life, the problem of this subject-object, or **inside-**outside, has a great meaning.

   (2) When the industrial revolution came up to **our** world, men started to work outside of their home and earn money from the **firm.**

   (3) Even when I go to the movie, I have to see advertising **of** restaurant, department store, clothes, and etc. before the movie starts.

   (4) When I was in eighth grade, I was a bad student.

Coding: circle **when** and put a '4"   above it.

5. **Extraposition** with sentence (or main-clause)-initial **it**

This category may be rather inferential in learners' production, but generally, extraposition refers to the movement of a 'heavy^n Subject such as a **that-clause** or infinitival Subject to postverbal position, and in its place the **dummy** pronoun **it** appears in pre-verbal position.

   (1) So, it is very hard to say that "What is the best way to study?'

   (2) Moreover, it is necessary to **recognize** problems exists in families today in order **to** explain the ideal role of a wife.

   (3) It is considered not manly if he cares much about his children or family in Japan.

   (4) It is hoped to develop the unfamiliar kinds of fishes.

Coding: underline the **it is (adj)** expression and above **it,** code it with **"5".**

Note: Some of the signals of this category are the phrases: "**it** is necessary/advisable/important/possible {for . . . to, that. . . }".

*Extraposition is described linguistically as follows (Baker 1978:43):

$$X - \underset{NP}{[S]} - Y$$
$$1 \quad 2 \quad 3$$
$$\Longrightarrow \quad 1, \ \underline{it}, \ 3 + 2 \ \text{(optional)}$$

This transformation is exemplified in the following a-b pairs:

a. That Fred won a prize amazed me.
b. It amazed me that Fred won a prize.
a. That Fred might win was believed by John.
b. It was believed by John that Fred might win.
a. For the Red Sox to lose now would annoy the reporters.
b. It would annoy the reporters for the Red Sox to lose now.

'These pairs of sentences illustrate a general regularity Of English: whenever we have a sentence whose subject is a sentential complement, there is a corresponding sentence in which <u>it</u> appears as the subject and the complement appears at the end." (Baker 1978:143)

6. True passives

These are distinguished from pseudo-passives, such as are found in 7a-8, below. They are not "true" in the sense that they are targetlike, however. Just to be conservative, we will count only those constructions with overt BE (not HAVE) plus the past participle (+/- -en); the <u>by</u> phrase need not appear.

(1) He was decided to be killed (2 passives).

(2) Specially in the developing countries in Africa and East Asia, the prominent tendency overpopulated is seen.

(3) After the principle of Malthus, the problem of population has been thought to be very important.

(4) A husband is expected to maintain his pride as a leader of his family.

151

<u>Coding</u>: Underline the passive construction and code it with "6".

Note that here we are testing for application of the following PASSIVE rule (Baker 1978:102-03):

$$NP - Aux - V - NP - X$$
$$1 \qquad 2 \quad 3 \qquad 4 \quad 5$$
$$\Longrightarrow \quad 4, \ 2, \ \underline{be} + \underline{en} + 3, \ 0, \ 5 + \underline{by} + 1 \text{ (optional)}$$

as in:

a. Sally amused the teacher.
b. The teacher was amused by Sally,

7.a. Topic-Comment: NP-∅-VP

This category includes Topics and "heavy Subjects", namely, sentences which serve as Subjects, but generally which do not reflect targetlike usage. Infinitives and gerunds serving as Subjects in a targetlike way are classed in 7.f. In some cases, the constructions will not have a verb, but rather, just a predicate (e.g., adjective). Thus they are NSs (non-syntactic units) as opposed to T-units, but they may be the clearest cases of this Topic-Comment construction. Because this category is, strictly speaking, NP-∅-VP, the initial phrase should be an NP (or a sentential NP). In English, we have topicalizing expressions for NPs such as "As for x," or "Concerning y," and these can be included in the NP category; "as to NP" will be treated in the same way as "as for NP."

With other expressions, such as prepositional phrases, though, we had better code the structure as a slightly different type; thus, we will code PP-∅-VP as 7aa, by convention. Below are examples of 7a:

(1) Take good physical care of themselves is very important (heavy Subject).

(2) Women go out of home and look for a job are usual today.

(3) "A life is compared to book, a wise read it carefully but a fool just read it unconcentrately." said by German priest, Lee.

(4) A man chooses a wife is a man's business.

(5) Of course his honest or not, self confident or not, genteer manners or not, is very important.

(6) "Those who sincerely wish to study should be given the first chance for college entrance, regardless of sex." is my opinion.

(7) As a wife in a family acts a very important role.

(8) Most of food which is served in such restaurant have cooked already.

(9) Chiang's food must make in the kitchen of the restaurant but Marty's food could make in his house.

Note that (8) and (9) look like they are attempts at passives; but in the literature, they are singled out as Topic-Comment constructions in which the Subject has been deleted (therefore, these are the "pseudo-passives".) Notice also that many times the "heavy Subject" or Topic is a quotation. It is difficult to define Topics in an operationalizable way (cf. Li and Thompson 1976), but if you find that the initial NP doesn't make sense as its Subject (e.g., it is clearly ungrammatical), although the predicate is a comment about it (e.g., #7 above), it is probably the Topic. More obvious cases are where there is no verb between the Topic NP and the Comment about it. (e.g., My house warm).

<u>Coding</u>: underline the TOPIC construction and code it as 7a.

For inter-rater reliablility purposes, ignore NP fragments of the type: "<u>For example the dogs</u>. There are many dogs at my house.' These will be analyzed as special kinds of Topic-Comments, but not here. Also, ignore discourse markers such as "At first," "Second," and "Last."

Again, note that there will be very similar types of constructions in which the initial element (TOPIC) is not an NP, but is a PP. This is an obvious attempt at trying to express the Topic case. (e.g., "<u>For books</u>, like novels best"); code these as 7aa.

7.b.   Topic–Comment **NP-NP-VP**

These are left–dislocations and double Subject constructions. Often they are introduced with the expressions **"As for"** or **"Concerning"**. If the Topic is not an NP in the usual type of Topic construction, but rather, a prepositional phrase which serves the same function (e.g., "<u>For books</u>, I like novels best"), code it as **7bb.**

(1) A lot of people, they know how to take good physical care by themselves.

(2) Homework, it is not a good thing.

(3) Older person, Katsu changed my life significantly.

(4) We, women including me are expected to be good mothers. (Notice the second NP has no **comma** following it; imagine the reverse: women including me, we are expected to be good mothers, which is an even clearer case).

(5) As for books, I like novels best.

<u>Coding</u>:   underline the construction (i.e. TOPIC) and code it 7b.

7.c.   Existential <u>there</u>-NP-VP

This is a type of construction with an existential plus a verb phrase with what appears to be a relative clause without a relative pronoun (also considered a serial verb).

(1) There are a small amount of people get married in their teenage.

(2) **A** man will be old, there are many things will be different.

(3) During the past 100 years, there were hundred of wars happened all around the world.

<u>Coding</u>:   Underline the <u>there</u> construction and code it 7c.

Note that <u>there</u> is meant to be an expletive, and not an adverbial. This transformation in English is generated by the following "THERE INSERTION" rule (Baker **1978:159**):

$$NP \text{ - } Aux \text{ - } \underline{be} \text{ - } Prep\ P$$
$$1 \qquad 2 \qquad 3 \qquad 4$$
$$==> \quad \underline{there}, \ 2 \ t \ 1, \ 4 \ (optional)$$

as in

    a.   **A** fly **is** in the **soup.**
    b.   There is a fly in the **soup.**

7.d.   Existential <u>there</u>-**BE** NP–to VP

    This is another existential construction with a different kind of attempt made **to** produce (or avoid?) a relative clause: an infinitive.

    (1)   There is no way to give some disadvantage women in getting into college.

    (2)   There are many elements to maintain a successful marriage.

    (3)   There is already enough people to populate **the earth.**

<u>Coding</u>:   Underline the <u>there</u> construction and code it **7d.**

7.e.   Existential <u>there</u>-**BE** NP–Relative Clause

    Finally, a (more–or–less) well–formed relative clause is made, that is, the relative pronoun is supplied where necessary.

    (1)   There are **alot** of people who thinks that way.

    (2)   There are many problems that can make marriage unsuccessful.

    (3)   **B**esides love, there are a lots of living problems you have to face.

    (4)   **There** are lots of ways that we can do to avoid this problem.

<u>Coding</u>:   Underline the construction and code it **7e.**

7'. Miscellaneous <u>there</u> constructions

> Because there is no category that includes <u>there</u> constructions without relative clauses (or attempted ones), or which includes other kinds of <u>there</u> **constructions,** ,this will be the <u>there</u> miscellaneous category.

(1) Therefore, we could say that there have good families there will have stable developing country.

(2) There is a picnic this weekend.

<u>Coding</u>: Underline the construction and code it **7'.**

7.f. Syntacticized **Topic/Subject** ‐ Predicate

This is the usual type of English **Subject-Predicate** construction. Included as Subjects are nominal **NPs,** gerunds, infinitives. The difference between the infinitive in **(2)** below and that in 7.a. **(1)** is that in this group, the **NPs** are basically **well-formed** and they are not strictly **"sentential"** Subjects. <u>We will not code all Subjects here, but just those with infinitival or gerundive Subjects. For this category, there must be</u> Subject-verb <u>agreement.</u>

(1) Having alot of money **isn't** important to me.

(2) To stick to those important rules isn't a hard job for me.

<u>Coding</u>: underline the Subject and code it **7f.**

8. Subject‐Verb Agreement

All Subject-verb constructions will be coded for agreement (or lack thereof); **"+"** =agreement; **"-"** =non-agreement*

(1) My father he go to work six **o'clock** [**-** AGR]
(2) Some people, he has a.lot of problem [+ **AGR**]
(3) My sister says she has **alot** of work to do
   [+ AGR; **+ AGR**]
(4) People has lots of problems my country
   [**- AGR**]

<u>Coding</u>: above each verb, put a "+" or a "-" to signify whether there is Subject–verb agreement. Disregard agreement in terms of tense. Only code auxiliaries which can carry number agreement (**e.g.**, <u>is</u>, <u>do</u>, <u>want</u>, <u>has</u>) and other verbs which are in the present tense; that is, ignore infinitives, and **if** you get any peculiar cases of "**he** <u>wills</u>" or "He wants **(+agr)** to <u>takes</u> (infinitive) a bath," just circle them as exceptional cases.

9.   Subject PRO–drop

This only considers those cases where a Subject is required in English surface structure. Supplied Subjects will not be tallied, only those that are erroneously omitted will be counted here. Pro–drop will be coded as "**PD**".

(1)   Depends on what time of year you go. **(PD)**
(2)   Because he went to school early, came home early. **(PD)**
(3)   She went home and took a nap (OKAY; not PD).
(4)   He said that to wants to go home today **(PD)**

<u>Coding</u>: place a "**PD**" above the missing Subject.

10.   Serial Verbs

Any serial verbs which are not counted in **7.c.**, will be counted here. Recall that **7.c.** constructions are like:

There <u>are</u> many people <u>have</u> a good time in Hawaii.

Basically, more than one main verb occurs in this sentence, therefore **it** is counted as a serial verb **(SV)** construction. Tally any <u>other</u> serial verb constructions (**i.e.**, not including these existential types) in this category. If the verbs are separated by commas, do not count them. Examples are:

(1)   Wife should go home take care of baby do housework **(SV=3)**

(2)   Alot of men, they **cook** dinner do the dishes (**SV=2**)

<u>Coding</u>: Underline the serial verb string and above it put "**SV**" and the number of concatenated verbs after **it.**

157

Note that some cases of serial verbs, as in those in 7a and 7c can be considered Topics since as they appear they are a kind of "heavy Subject."

Li and Thompson (1981:620) identify serial verbs (e.g., in Mandarin) as follows: "there are at least two verbs (and various numbers of noun phrases) that are concatenated to express one overall event or state of affairs." The string is:

(NP) V (NP) (NP) V (NP)

11. Topic Chains

These will be tallied separately.

Table A

Summary of Frequency Statistics for all Dependent Measures Broken Down by Ll, Proficiency, and Ll x Proficiency

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
LOC 2a (In)- LOC-VP	Entire Group	7.0	0.03	0.20	210	
	By Ll/ J	3.0	0.03	0.17	105	
	M	4.0	0.04	0.24	105	
	By PROF/low	3.0	0.04	0.20	70	
	medium	4.0	0.06	0.29	70	
	high	0.0	0.00	0.00	70	
	Ll J PROF low	2.0	0.06	0.26	35	
	medium	1.0	0.03	0.17	35	
	high	0.0	0.00	0.00	35	
	Ll M PROF low	1.0	0.03	0.17	35	
	medium	3.0	0.09	0.37	35	
	high	0.0	0.00	0.00	35	
LOC 2b In + LOC NP-VP	Entire Grp	82.0	0.39	0.72	210	
	By Ll/ J	43.0	0.41	0.78	105	
	M	39.0	0.37	0.65	105	
	By PROF/low	33.0	0.47	0.78	70	
	medium	21.0	0.30	0.49	70	
	high	28.0	0.40	0.84	70	
	Ll J PROF low	15.0	0.43	0.70	35	
	medium	9.0	0.26	0.51	35	
	high	19.0	0.54	1.04	35	
	Ll M PROF low	18.0	0.51	0.85	35	
	medium	12.0	0.34	0.48	35	
	high	9.0	0.26	0.56	35	

Depend. Measure	Independent Variables		Sum	Mean	Standard Deviation	Cases	Signif
LOC 2bb	Entire Grp		23.0	0.11	0.33	210	
In + TEMP							
NP + VP	By Ll/	J	7.0	0.07	0.25	105	
		M	16.0	0.15	0.39	105	
	By PROF/low		9.0	0.13	0.34	70	
	medium		8.0	0.11	0.36	70	
	high		6.0	0.09	0.28	70	
	Ll J						
	PROF low		2.0	0.06	0.25	35	
	medium		3.0	0.09	0.28	35	
	high		2.0	0.06	0.24	35	
	Ll M						
	PROF low		7.0	0.20	0.41	35	
	medium		5.0	0.14	0.43	35	
	high		4.0	0.11	0.32	35	
Tot. 2B	Entire Grp		105.0	0.50	0.83	210	
2b + 2bb							
(In-	By Ll/	J	50.0	0.48	0.84	105	
NP-NP-VP)		M	55.0	0.53	0.82	105	
	By PROF/low		42.0	0.60	0.91	70	
	medium		29.0	0.41	0.67	70	
	high		34.0	0.49	0.90	70	
	Ll J						
	PROF low		17.0	0.49	0.78	35	
	medium		12.0	0.34	0.59	35	
	high		21.0	0.60	1-09	35	
	Ll M						
	PROF low		25.0	0.71	1.02	35	
	medium		17.0	0.49	0.74	35	
	high		13.0	0.37	0.65	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
LOC 2c	Entire Grp	22.0	0.10	0.35	210	
There +						
VP +	By Ll/ J	11.0	0.10	0.34	105	
in LOC	M	11.0	0.10	0.36	105	
	By PROF/low	10.0	0.14	0.43	70	
	medium	9.0	0.13	0.38	70	
	high	3.0	0.04	0.20	70	
	Ll J					
	PROF low	6.0	0.17	0.45	35	
	medium	3.0	0.09	0.28	35	
	high	2.0	0.06	0.24	35	
	Ll M					
	PROF low	4.0	0.11	0.40	35	
	medium	6.0	0.17	0.45	35	
	high	1.0	0.03	0.17	35	
LOC 2	Entire Grp	111.0	0.53	0.84	210	
Tot. LOC						
2a+2b+2c	By Ll/ J	57.0	0.54	0.88	105	
	M	54.0	0.51	0.80	105	
	By PROF/low	46.0	0.66	0.92	70	
	medium	34.0	0.49	0.74	70	
	high	31.0	0.44	0.85	70	
	Ll J					
	PROF low	23.0	0.66	0.94	35	
	medium	13.0	0.37	0.60	35	
	high	21.0	0.60	1.03	35	
	Ll M					
	PROF low	23.0	0.66	0.91	35	
	medium	21.0	0.60	0.85	35	
	high	10.0	0.29	0.57	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
If (Sentence-Initial)	Entire Grp	93.0	0.44	0.78	210	
	By Ll/ J	52.0	0.50	0.82	105	
	M	41.0	0.39	0.73	105	
	By PROF/low	34.0	0.49	0.88	70	
	medium	37.0	0.53	0.81	70	
	high	22.0	0.31	0.60	70	
	L1 J PROF low	22.0	0.63	1.00	35	
	medium	20.0	0.57	0.81	35	
	high	10.0	0.29	0.57	35	
	L1 M PROF low	12.0	0.34	0.73	35	
	medium	17.0	0.49	0.82	35	
	high	12.0	0.34	0.64	. 35	
When (Sentence-Initial)	Entire Grp	55.0	0.26	0.56	210	
	By Ll/ J	29.0	0.28	0.58	105	
	M	26.0	0.25	0.53	105	
	By PROF/low	22.0	0.31	0.71	70	
	medium	19.0	0.27	0.48	70	
	high	14.0	0.20	0.44	70	
	L1 J PROF low	13.0	0.37	0.77	35	
	medium	9.0	0.26	0.44	35	
	high	7.0	0.20	0.47	35	
	L1 M PROF low	9.0	0.26	0.66	35	
	medium	10.0	0.29	0.52	35	
	high	7.0	0.20	0.41	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Total If+When	Entire Grp	148.0	0.70	0.92	210	
	By Ll/ J	81.0	0.77	0.96	105	
	M	67.0	0.64	0.88	105	
	By PROF/low	56.0	0.80	1.08	70	
	medium	56.0	0.80	0.96	70	
	high	36.0	0.51	0.65	70	
	Ll J PROF low	35.0	1.00	1.21	35	
	medium	29.0	0.83	0.89	35	
	high	17.0	0.49	0.66	35	
	Ll M PROF low	21.0	0.60	0.91	35	
	medium	27.0	0.77	1.03	35	
	high	19.0	0.54	0.66	35	
It (Extra-position)	Entire Grp	105.0	0.50	0.74	210	
	By Ll/ J	61.0	0.58	0.76	105	
	M	44.0	0.42	0.72	105	
	By PROF/low	32.0	0.46	0.74	70	
	medium	39.0	0.56	0.85	70	
	high	34.0	0.49	0.63	70	
	Ll J PROF low	17.0	0.49	0.70	35	
	medium	21.0	0.60	0.85	35	
	high	23.0	0.66	0.73	35	
	Ll M PROF low	15.0	0.43	0.78	35	
	medium	18.0	0.51	0.85	35	
	high	11.0	0.31	0.47	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Passive	Entire Grp	234.0	1.11	1.33	210	
	By Ll/ J	145.0	1.38	1.58	105	***
	M	89.0	0.85	1.02	105	p < .005
	By PROF/low	58.0	0.83	1.09	70	
	medium	74.0	1.06	1.19	70	p < .05
	high	102.0	1.46	1.66	70	*
	Ll J PROF low	32.0	0.91	1.29	35	*
	medium	40.0	1.14	1.29	35	p < .05
	high	73.0	2.09	1.88	35	*
	Ll M PROF low	26.0	0.75	0.85	35	
	medium	34.0	0.97	1.10	35	
	high	29.0	0.83	1.10	35	
TC 7a NP-Ø-VP	Entire Grp	32.0	0.15	0.46	210	
	By Ll/ J	11.0	0.10	0.34	105	
	M	21.0	0.20	0.56	105	
	By PROF/low	13.0	0.19	0.46	70	
	medium	13.0	0.19	0.60	70	
	high	6.0	0.09	0.28	70	
	Ll J PROF low	5.0	0.14	0.43	35	
	medium	4.0	0.11	0.32	35	
	high	2.0	0.06	0.24	35	
	Ll M PROF low	8.0	0.23	0.49	35	
	medium	9.0	0.26	0.78	35	
	high	4.0	0.11	0.32	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
TC 7b NP-NP-VP	Entire Grp	63.0	0.30	0.60	210	
	By Ll/ J	28.0	0.27	0.56	105	
	M	35.0	0.33	0.63	105	
	By PROF/low	24.0	0.34	0.61	70	
	medium	15.0	0.21	0.48	70	
	high	24.0	0.34	0.68	70	
	Ll J PROF low	12.0	0.34	0.64	35	
	medium	4.0	0.11	0.32	35	
	high	12.0	0.34	0.64	35	
	Ll M PROF low	12.0	0.34	0.59	35	
	medium	11.0	0.31	0.58	35	
	high	12.0	0.34	0.73	35	
TC 7bb PP-NP-VP	Entire Grp	40.0	0.19	0.52	210	
	By Ll/ J	14.0	0.13	0.42	105	
	M	26.0	0.25	0.60	105	
	By PROF/~	12.0	0.17	0.42	70	
	medium	13.0	0.19	0.49	70	
	high	15.0	0.21	0.63	70	
	Ll J PROF low	5.0	0.14	0.43	35	
	medium	5.0	0.14	0.49	35	
	high	4.0	0.11	0.3 2	35	
	Ll M PROF low	7.0	0.20	0.41	35	
	medium	8.0	0.23	0.49	35	
	high	11.0	0.31	0.83	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Tot. 7B TC 7b+ 7bb (PP/NP- NP-VP)	Entire Grp	103.0	0.49	0.85	210	
	By **Ll/** J	42.0	0.40	0.72	105	
	M	61.0	0.58	0.96	105	
	By **PROF**/low	36.0	0.51	0.83	70	
	medium	28.0	0.40	0.73	70	
	high	39.0	0.56	0.97	70	
	Ll J PROF low	17.0	0.49	0.24	35	
	medium	9.0	0.26	0.37	35	
	high	16.0	0.46	0.24	35	
	Ll M PROF low	19.0	0.54	0.53	35	
	medium	19.0	0.54	0.47	35	
	high	23.0	0.66	0.40	35	
TC 7c There + VP-VP	Entire Group	8.0	0.04	0.22	210	
	By Ll/ J	3.0	0.03	0.17	105	
	M	5.0	0.05	0.26	105	
	By PROF/low	4.0	0.06	0.23	70	
	medium	2.0	0.03	0.17	70	
	high	2.0	0.03	0.24	70	
	Ll J PROF low	2.0	0.06	0.24	35	
	medium	1.0	0.03	0.17	35	
	high	0.0	0.00	0.00	35	
	Ll M PROF low	2.0	0.06	0.24	35	
	medium	1.0	0.03	0.17	35	
	high	2.0	0.06	0.34	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
TC 7d There+VP to VP	Entire Grp	10.0	0.48	0.27	210	
	By L1/ J	3.0	0.03	0.17	105	
	M	7.0	0.07	0.35	105	
	By PROF/low	3.0	0.04	0.27	70	
	medium	2.0	0.03	0.24	70	
	high	5.0	0.07	0.31	70	
	L1 J PROF low	0.0	0.00	0.00	35	
	medium	0.0	0.00	0.00	35	
	high	3.0	0.09	0.28	35	
	L1 M PROF low	3.0	0.09	0.37	35	
	medium	2.0	0.06	0.34	35	
	high	2.0	0.06	0.34	35	
TC 7e There+VP +Rel Cl	Entire Grp	14.0	0.07	0.27	210	
	By L1/ J	8.0	0.08	0.27	105	
	M	6.0	0.06	0.27	105	
	By PROF/low	2.0	0.03	0.17	70	
	medium	6.0	0.09	0.28	70	
	high	6.0	0.09	0.33	70	
	L1 J PROF low	1.0	0.03	0.17	35	
	medium	4.0	0.11	0.32	35	
	high	3.0	0.09	0.28	35	
	L1 M PROF low	1.0	0.03	0.17	35	
	medium	2.0	0.06	0.24	35	
	high	3.0	0.09	0.37	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
TC 7f Infin/ Gerund + VP (+AGR)	Entire Grp	37.0	0.18	0.47	210	
	By Ll/ J	16.0	0.15	0.43	105	
	M	21.0	0.20	0.51	105	
	By PROF/low	7.0	0.10	0.30	70	
	medium	14.0	0.20	0.44	70	
	high	16.0	0.23	0.62	70	
	Ll J PROF low	4.0	0.11	0.32	35	
	medium	7.0	0.20	0.41	35	
	high	5.0	0.14	0.55	35	
	Ll M PROF low	3.0	0.09	0.28	35	
	medium	7.0	0.20	0.47	35	
	high	11.0	0.31	0.68	35	
TC 7' There (exist) misc.	Entire Grp	76.0	0.36	0.67	210	
	By Ll/ J	39.0	0.37	0.64	105	
	M	37.0	0.35	0.69	· 105	
	By PROF/low	19.0	0.27	0.56	70	
	medium	32.0	0.46	0.79	70	
	high	25.0	0.36	0.61	70	
	Ll J PROF low	12.0	0.34	0.64	35	
	medium	12.0	0.34	0.64	35	
	high	15.0	0.43	0.65	35	
	Ll M PROF low	7.0	0.20	0.47	35	
	medium	20.0	0.57	0.92	35	
	high	10.0	0.29	0.57	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Tot. TC7 7a-7f	Entire Grp	204.0	0.97	1.24	210	
	By L1/ J	83.0	0.79	1.01	105	*
	M	121.00	1.15	1.43	105	p < .05
	By PROF/low	65.0	0.93	1.03	70	
	medium	65.0	0.93	1.39	70	
	high	74.0	1.06	1.32	70	
	L1 J PROF low	29.0	0.83	0.98	35	
	medium	25.0	0.71	0.93	35	
	high	29.0	0.83	1.12	35	
	L1 M PROF low	36.0	1.03	1.07	35	
	medium	40.0	1.14	1.72	35	
	high	45.0	1.29	1.47	35	
Total There (7c+7d+ 7e+7f+ 7'+2c)	Entire Grp	108.0	0.51	0.84	210	
	By L1/ J	53.0	0.50	0.77	105	
	M	55.0	0.52	0.91	105	
	By PROF/~	28.0	0.40	0.79	70	
	medium	42.0	0.60	0.91	70	
	high	38.0	0.54	0.83	70	
	L1 J PROF low	15.0	0.43	0.70	35	
	medium	17.0	0.49	0.74	35	
	high	21.0	0.60	0.88	35	
	L1 M PROF low	13.0	0.37	0.88	35	
	medium	25.0	0.71	1.05	35	
	high	17.0	0.49	0.78	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Total Dummy Subjects (It + There)	Entire Grp	213.0	1.01	1.18	210	
	By Ll/ J	114.0	1.09	1.19	105	
	M	99.0	0.94	1.17	105	
	By PROF/low	60.0	0.86	1.04	70	
	medium	81.0	1.16	1.34	70	
	high	72.0	1.03	1.14	70	
	Ll J PROF low	32.0	0.91	1.04	35	
	medium	38.0	1.09	1.31	35	
	high	44.0	1.26	1.22	35	
	Ll M PROF low	28.0	0.80	1.05	35	
	medium	43.0	1.23	1.37	35	
	high	28.0	0.80	1.02	35	
AGR 1 Verbs with S-V AGR (raw scores)	Entire Grp	2243.0	10.68	4.05	210	
	By Ll/ J	1164.0	11.09	3.91	105	
	M	1079.0	10.28	4.17	105	
	By PROF/low	682.0	9.74	3.84	70	
	med	780.0	11.14	3.75	70	
	high	781.0	11.16	4.44	70	
	Ll J PROF low	359.0	10.26	3.69	35	
	med	399.0	11.40	3.77	35	
	high	408.0	11.66	4.22	35	
	Ll M PROF low	325.0	9.29	3.98	35	
	med	381.0	10.89	3.76	35	
	high	373.0	10.66	4.65	35	

Depend. Measure	Independent Variables		Sum	Mean	Standard Deviation	Cases	Signif
AGR 2 Verbs with IL AGR (no plural with QUANT + NP) (raw scores)	Entire Grp		30.0	0.14	0.44	210	
	By Ll/	J	14.0	0.13	0.42	105	
		M	16.0	0.15	0.46	105	
	By PROF/	low	15.0	0.21	0.61	70	
		medium	9.0	0.13	0.34	70	
		high	6.0	0.09	0.28	70	
	Ll J PROF	low	7.0	0.20	0.58	35	
		medium	3.0	0.09	0.28	35	
		high	4.0	0.11	0.32	35	
	Ll M PROF	low	8.0	0.23	0.65	35	
		medium	6.0	0.17	0.38	35	
		high	2.0	0.06	0.24	35	
AGR 3 Verbs w/o AGR (raw scores)	Entire Grp		186.0	0.89	1.17	210	
	By Ll/	J	75.0	0.71	0.99	105	* $p < .05$
		M	111.0	1.06	1.32	105	
	By PROF/	low	76.0	1.09	1.39	70	
		medium	66.0	0.94	1.18	70	
		high	44.0	0.63	0.87	70	
	Ll J PROF	low	28.0	0.80	1.11	35	
		medium	23.0	0.66	0.84	35	
		high	24.0	0.69	1.02	35	
	Ll M PROF	low	48.0	1.37	1.59	35	
		medium	43.0	1.23	1.40	35	
		high	20.0	0.57	0.70	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
% AGR 1 Verbs with AGR divided by AGR1+2+3	Entire Grp		0.91	0.12	210	
	By L1/ J		0.93	0.09	105	*
	M		0.89	0.12	105	p < .05
	By PROF/ low		0.89	0.15	70	
	medium		0.91	0.11	70	
	high		0.94	0.09	70	
	L1 J PROF low		0.91	0.11	35	
	medium		0.94	0.07	35	
	high		0.94	0.10	35	
	L1 M PROF low		0.86	0.17	35	
	medium		0.88	0.13	35	
	high		0-93	0.09	35	
% AGR 2 Verbs with IL AGR divided by AGR1+2+3	Entire Grp		0.01	0.04	210	
	By L1/ J		0.01	0.04	105	
	M		0.01	0.03	105	
	By PROF/low		0.02	0.05	70	
	medium		0.01	0.03	70	
	high		0.01	0.03	70	
	L1 J PROF low		0.02	0.05	35	
	medium		0.01	0.02	35	
	high		0.01	0.04	35	
	L1 M PROF low		0.01	0.04	35	
	medium		0.01	0.03	35	
	high		0.01	0.03	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
% AGR 3 (verbs	Entire Grp		0.08	0.08	210	
w/o AGR	By L1/ J		0.06	0.08	105	**
divided	M		0.10	0.13	105	p < .01
by						
AGR1+2+3	By PROF/low		0.10	0.14	70	
	medium		0.08	0.10	70	
	high		0.06	0.08	70	
	L1 J PROF low		0.07	0.10	35	
	medium		0.05	0.06	35	
	high		0.05	0.08	35	
	L1 M PROF low		0.13	0.17	35	
	medium		0.10	0.13	35	
	high		0.06	0.08	35	
PRO-Drop (Subj)	Entire Grp	58.0	0.28	0.79	210	
	By L1/ J	14.0	0.13	0.39	105	**
	M	44.0	0.42	1.05	105	p < .01
	By PROF/low	24.0	0.34	0.78	70	
	medium	23.0	0.33	1.06	70	
	high	11.0	0.16	0.44	70	
	L1 J PROF low	10.0	0.26	0.56	35	
	medium	4.0	0.11	0.32	35	
	high	1.0	0.03	0.17	35	
	L1 M PROF low	15.0	0.43	0.95	35	
	medium	19.0	0.54	1.44	35	
	high	10.0	0.29	0.57	35	

Depend. Measure	Independent Variables	Sum	Mean	Standard Deviation	Cases	Signif
Serial Verbs (not 7c)	Entire Grp	21.0	0.10	0.33	210	
	By Ll/ J	4.0	0.39	1-06	105	**
	M	17.0	0.16	0.40	105	p < .01
	By P R O F / ~ ~	9.0	0.13	0.38	70	
	medium	8.0	0.11	0.36	70	
	high	4.0	0-06	0.23	70	
	Ll J PROF low	0.0	0.00	0.00	35	
	medium	2.0	0.06	0.34	35	
	high	2.0	0.06	0.24	35	
	Ll M PROF low	9.0	0.26	0.51	35	
	medium	6.0	0.17	0.38	35	
	high	2.0	0.06	0.24	35	
Total Serial Verbs SV+7c)	Entire Grp .	29.0	0.14	0.39	210	
	By Ll/ J	7.0	0.07	0.29	105	**
	M	22.0	0.21	0.47	105	p < .01
	By PROF/low	13.0	0.19	0.43	70	
	medium	10.0	0.14	0.43	70	
	high	6.0	0.09	0.33	70	
	Ll J PROF low	2.0	0.06	0.24	35	
	medium	3.0	0.09	0.37	35	
	high	2.0	0.06	0.24	35	
	Ll M PROF low	11.0	0.31	0.53	35	
	medium	7.0	0.20	0.47	35	
	high	4.0	0.11	0.40	35	

APPENDIX C

EXCERPTS FROM COMPOSITIONS

Japanese

Level 1

(S5)
1. Today's world, there are two features.
2. However₁ there is the dangerous of advertising.

(S33)
3. It have been become very convenient our daily life.
4. Recently, Japanese was called 'Economic Animal".
5. For example, if our human beings use TV and radio, we can understand that what happen today all of the world.
6. If we use car, we can go any place in the country.
7. If we use washing machine₁ we can cut off the time instead of hand washing.

(S36)
8. In junior high school₁ however₁ there is obvious discrimination in Japan.
9. Most high schools₁ in Japan are made for girls and boys separately.
10. And 2/3 of high school students₁ in Japan began to attend the universities.
11. Many people said that it was easier to educate them according to the character of different sex.
12. Therefore₁ in the girls school₁ the time of political science class of physics class is reduced and they are forced to learn cooking₁ sewing₈ washing closes₁ etc.
13. On the other hand, boys at high schools₁ they learn more scientific things and they are encouraged to study hard for the university.

(S41)
14. In Japan, men are certainly dominant because of history.
15. In Tokugawa Era, Eeyasu Tokugawa was interested in studying Confucius in China.
16. If the men has a concumbinage, the wife has to be patient₈ in order to keep the family.
17. In 1980s the time has been changed.
18. I believe, it is important to have women's liberation.

19. But in Japan, they still have distinguish between men and women.

(S54)
20. The hardest period in my life, it's I think, the time I graduated college.
21. The teacher, he is a science teacher, influenced me.
22. But when I really became a school teacher, I understand it's very hard work.
23. Moreover, there are a lot of students. All of them has each character.

J Level 2

(S12)
24. Then some politicians emphasized that to make Japan international country English must be used in education.
25. If their opinion won, now maybe I can speak English very well,-write English very well.
26. So now we are educated in our own language.
27. But when it began there were many difficult, problem.
28. Thesis, words or something like that, had to translate into Japanese.
29. In Hawaii always blow wind named 'Traffic Wind".
30. In Japanese we call it "Boueki fu".
31. They are difficult to explain what they mean.

(S43)
32. Father Burt Gramelspacher. He gave a great inluence to my life.
33. Most of trainees are sent from their company to study there for three monthes.
34. And on his room's door, there was a sign says "Dirty old man here".

(S72)
35. It can be said "human," when people have these three independence.
36. It cannot be said 'men should be given the first choice for college entrance? for they are the chief wage earners.'

(S87)
37. To attend an American university gives me a great benefit in my life.
38. In homeland, we have to read or sometimes write in English.

176

39. Of course we are obliged to do in English in America however.
40. I think that it is desirable to attend an American school after some years English training in homeland.
41. We, foreigners can easily apply American universities, study in detail also.

J Level 3

(S6)
42. If you want to visit Niagara Falls, watch television and you can see the gigantic view with 4 channel sound.
43. If you want to see a movie, watch television and you can see it for free.
44. I have to admit that there are a few good points of watching TV such as educational TV.

(S37)
45. If he/she goes as a sightseener and look around for a week or so, I recommend him/her to visit Tokyo, Osaka, Kyoto, and Nara.
46. It is not exaggerated to say that all the people from abroad are surprised to see the numbers of people.
47. There are so many people that it is hard to find a good house or apartment to live.
48. To Osaka a visitor will have nearly the same expression as he has on Tokyo.
49. As for Kyoto and Nara, there are alot of books and advertisements on them.
50. So sometimes there are visitors who know more than Japanese about them.

(S67)
51. After the principle of Malthus, the problem of population has been thought to be very important.
52. I think that the problem is very important and should be resolved as soon as possible.
53. Taking the case of food, I think that the further growth of production is hoped rather than on the land.
54. We eat only small kinds of fishes, but there are many kinds of fishes in the sea of the world.
55. It is hoped to develop the unfamiliar kinds of fishes.
56. Finally, considering the birth control, I don't know what is the best way of it because I am a singular.

57. I hear there is no perfect way to avoid to become pregnant.

(S74)
58. (This is natural for Americans). But as for me, it was sensational, since in Japan I had to study English as compulsory.
59. When I *was* in Japan after becoming in English, I kept on studying English and taught myself,

Japanese TCs (7a, 7b, 7bb)

Level 1

1. The main reason, of course, I want to study English (S3).
2. For teaching English in Japan, we will take advantage if we are graduated from university in America or England (S3).
3. For example, college life, how to study, what is the main point of them etc. (S3)
4. The fact that married couple will reduce having their children alot if the government will not pay them alot (S24).
5. Most high schools, in Japan are made for girls and boys separately (S36).
6. And 2/3 of high school students, in Japan began to attend the universities (S36).
7. On the other hand, boys at high schools, they learn more scientific things and they are encouraged to study hard for the university (S36).
8. The hardest period of my life, it's I think, the time I graduate college (S54).
9. The teacher, he is a science teacher, influenced me (S54).
10. From the viewpoint of a small country, we are able to find both advantage and disadvantage to be governed by a large country (S78).
11. From the viewpoint of a large country, we are also able to find both advantage and disadvantage to be governed by a large country (S78).
12. According to world history, the biggest war, World War I and World War II breaked after large increasing in the population (S89).
13. A wife as a woman, she have to obey her husband even if she knows that he is wrong (S99).
14. For people who are too busy to go shopping or living in the countryside far from the city, it's very convenient system (S102).

15. **And** the way **I** was (**,**) so different from the way the host parents treat their children **(S10)**.

16. **Thesis,** words or something like that had to translate into Japanese **(S12)**.

17. (The young **president,** John F. **Kennedy,** put emphasis on vitality and vigor.) On the **contrary,** the 71-year old **president,** Ronald Reagan wants the United States of America to be the good and old America **(S20)**.

18. As to the positive **aspects,** here are some **(S20)**.

19. . . . and the way they treated **me.** They **didn't** treat me the way **I** was expected **(S31)**.

20. During the period while **I** was getting used to the new **culture,** I call it "the hardest period in my life" **(S31)**.

21. To Osaka a visitor will have nearly the same expression as he has on Tokyo **(S37)**.

22. As for Kyoto and **Nara,** there are **alot** of books and advertisements on them **(S37)**.

23. **Actually, I** didn't like him because for **me,** he was a kind of cramming type **of** person who was always studying in order to get good grades **(S40)**.

24. No discrimination toward aged people in studying. What a liberalized idea **(S50)**.

25. In other **words,** for survival daily **necessities,** for **example, cleaning,** cooking, washing are very important **(S59)**.

26. After the principle of **Malthus,** the problem of population has been thought to be very important **(S67)**.

27. Taking the case of **food, I** think that the further growth of production in the sea is hoped rather than on the land **(S67)**.

28. (This **is** natural for **Americans).** But as for **me,** it was **sensational,** since in Japan **I** had to study English as compulsory.

Mandarin

Level 1

(S109)
1. But the water is polluted.
2. We drink it just like drink the poison.
3. The poison go into our body destroy our living cells.
4. How bout animals, don't they drink water?
5. When we breathe those air, then into our lungs.

(S120)
6. If the political situation is stead (I think this is the most important thing in China and I hope so) I'm sure that our country will develop quickly.
7. There will be a refrigerator in kicken.
8. It will be achieved.
9. If the political situation is firm, I hope that every people in our country will not be confined so much.

(S129)
10. It coursed a lot of fish, plants to be killed.
11. And even people themselves, they got strange sick.
12. But how to cut down the pollution immediately, it's really a big problem.
13. If we can't find a good method one day, the world will become a dirty dark place.
14. But if we consider the goal of our life, we just want to have a safe, comfortable, wealthy food place.

(S130)
15. In outcountry President is as our people.
16. There are six parts of our goverment that manerging the country.
17. There is no law give more power to govenet and limite the freedom of people.
18. So in Taiwan there is no strike.
19. And strike is no need.

(S138)
20. In the meantime, these case won't happen to an American family.
21. It comes another problem to me to overcome, it is the problem of normal courses.
22. It was very hard for me to catch up with the teachers at the beginning,

23. When I arrived at the airport, I immediately recognized the the sharp differences between a western and an oriental countries.

M Level 2

(S106)

24. There are many criteria to attack or defend this statement.
25. There are so many people they don't know why they want to be in college.
26. In respect of above main goals of education, one should know, there are nothing to do of the sex difference in the choice of who will be considered to be entranced in college.
27. The one who meet the requirement" or who has the potential to be educated for the goals is real one.

(S121)

28. As we know, there is a ten years break in Chinese education.
29. It is hardly to imagine that one can make good progress in his study in such a college.
30. There are plenty of scientists and professors in American college. They are working or solving most difficult problem.
31. In China, we have some good professional people.
32. China need advanced knowledge to perform her modernization.

(S155)

33. When I was nineteen years old, I finally got my father's permit to learn Kung Fu.
34. Chang, Tse-Tung -- my seefu's name, sixty years old, a man who was a very famous martial artist during his early years.
35. In Chinese, See means teacher.
36. He is not a kind of people who talk alot, but do a little.
37. Practicing Kung Fu is a hard work.
38. There were a few times, I almost tried to give up,
39. My master kept couraging me, asked me not to give up.
40. It has been eight years since I practiced Kung Fu.

(S185)

41. As an adult, the work, the family and the friends will influence one's life.
42. So wheater the attitude will go on possitive side

or negative side, depend upon on one's living environment.

43. Be more aware of what's happening, increase understanding toward itself, others, and to the situation.
44. Capable, well experience and have a better judgment.
45. Afraid to face the reality.
46. "Aging," negative feelings are involve.
47. Feel useless and powerless.
48. Experience the physical crisis, incapable at some degree.
49. From my points of view, aging is the process of feeling involve.
50. Is an emotional things.

M Level 3

(S111)
51. To increase our agricultural product and other natural sources -- we may call it our essential requirements to live.
52. such as synthetic fuel or energy that can be made by man.
53. And to migrate people on other planets is also a best way to solve the problem.
54. An old say, there is a will, there is a way.

(S145)
55. In a class those who are good at competition will get better and better in the school work.
56. This idea is reflected to the religion.
57. As a developing country, we now stick to our survival, the govenrment do every thing with extreme efforts.
58. Places like Hong Kong, Singapore, China and etc, there are so populated that some of them become lack of food.
59. If the population of the earth continues to increase at the present rate, there will be many country ends with the same problem as india did right now.
60. Birth control is used.
61. More than this limitation will have to pay a certain tax.
62. There also built up several birth control institute to give knowledge of birth control.
63. Secondly, every country should have their law that prevents outsiders from permanently resides in

64. Another point of view which I strongly against it is -- men are the chief wage earner.
65. And as for the women, having intelligence and talent as well as men, can also showing their efficiency on their work, and get a important and higher position in a company or in the government.
66. From only this point view, we can easyly to deny that men is not the only chief wage earners in this society.
67. But there is no connection that men should be given the first choice for college entrance.
68. Besides₁ in this country, we get enough colleges to maintain the people who want to be more educated.
69. As for the last point, I want to say yes, it is nature that taking care of children and housekeeping is the woman's responsibility.

(S175)
70. In some families it is not unusual that the woman is in fact the chief wage earner.
71. In a family, men and women should share and have the responsibility of their home and children.
72. As far as a family is concerned, men and women should share everything.
73. Having an education, can train a person to understand things better and to learn how to solve problems.

(S178)
74. In civilization, people, either woman or man, should have a higher education.
75. Although taking care of children is woman's talent, women should have the same opportunity to go to college and educate themselves.
76. It is very ridiculous to forbid woman to go to the college.
77. In many aspects, we can see many women are very capable in their career.
78. For example the Prime Minister of England. Clearly, her ability goes beyond many men.
79. One thing for sure, if hshe doesn't have the same choice as man, she would not be the Prime Minister.
80. If women have the same opportunity, they can become the chief wage-earners too.
81. The first choice for college entrance should not give to men.

Mandarin TCs (7a, 7b, 7bb)

<u>Level 1</u>

1. TV stars, singers are the person who the women and girls like (S131).
2. Especially the commercial, women buy anything that the TV says very good (S131).
3. For this problem, we had to change their opinions (S152).
4. The population increased too fast that made many countries support life (S152).
5. And for his opinion, he thinks as long as he can make enough money to support our family is enough (S180).
6. From the examination* students can really know how much they learn from the class (S183).
7. Now alot of Chinese, especially young people would stay home (S110).
8. For school, I have seen that the student of US are have so much freedom (S186).
9. "Women should concentrate on the home and children" is a unfair statement (S206).
10. For alot of Chinese, they though the more kids you have, they richer you are (S210).
11. Thoughts, television gave us lots of good, it gave us some bad idea too (S126).
12. In our television show always have program from Korea, America and other countries (S126).
13. How bout animals, don't they drink water (S109)?
14. And even people themselves, they got strange sick (S129).

<u>M</u> <u>Level 3 TCs</u>

15. As for the last point, I want to say *yes*, it is nature that taking care of children and doing housekeeping is the woman's job (S169).
16. As far as a family is concerned, men and women should share everything (S175).
17. As to the job itself, my capability and ability of handling my responsibilities was really more than enough (S200).
18. To the clevest students he will not tell them the exact answers, but encourage them to think themselves (S202).
19. To the dull ones, he will show them how to get to the right place quickly (S202).
20. More peace-loving, and more enjoying the sexual intercourse ⎯⎯ it makes more infants born (S133).

21. The improvement of living environment: we have more powerful protection *to* all kinds of diseases etc (S133).
22. An old say, there is a will, there is a way (S111).
23. From these films, we can feel the warmth of the human beings and find the value to live (S165).
24. How to maintain a family being happy all the time? I think it is a big thing and not easy (S182).
25. From the point of sociology, society looks just like rolling wheel (S190).

BIBLIOGRAPHY

Andersen, R. 1983. Transfer to somewhere. In S. Gass and L. Selinker (eds.), 177-201.

Baker, C. 1978. Introduction to aenerative transformational svntax. Englewood Cliffs, New Jersey: Prentice-Hall.

Barry, R. 1975. Topic in Chinese: An overlap of meaning, grammar, and discourse function. In R. Grossman, L. San, and T. Vance (eds.), Papers from the parasession on functionalism. Chicago Linguistic Society. 1-9.

Bates, E. 1976. Language and context: the acquisition of pragmatics. New York: Academic Press.

Bates, E., and B. MacWhinney. 1979. A functionalist approach to the acquisition of grammar. In E. Ochs and B. Schieffelin (eds.), Developmental praamatics. New York: Academic Press. 167-212

Bland, S. 1981. Topic-comment sentences in English. In Cornell Working Papers in Linguistics 2. 32-49.

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (ed.), 25-56.

Chaudron, C. 1983. Simplification of input, topic reinstatements and their effects on L2 learners' recognition and recall. TESOL Quarterly 17,3. 437-458.

Chen, P. 1984. A discourse analysis of third person zero anaphora in Chinese. Bloomington, Indiana: Indiana University Linguistics Club.

Chomsky, N. 1977. On &-movement. In P. Culicover, T. Wasow, and A. Akma-jian (eds.), Formal syntax. New York: Academic Press. 71-132.

Chomsky, N. 1981. Lectures on government and binding. Dordrecht: Foris.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20. 37-46.

Cole, P., W. Harbert, G. Hermon, and S. Sridhar. 1980.
    The acquisition of subjecthood. <u>Language</u> 56,4.
    719-743.

Corder, S. 1983. A role for the mother tongue. In S.
    Gass and L. Selinker (eds.), 85-97.

Curran, S. 1984. Transfer and topic prominence in Korean-
    English interlanguage. Unpublished ms. University
    of California at Los Angeles.

de Groot, C. 1981. On theme in functional grammar. In T.
    Hoekstra, H. van der Hulst, and M. Moortgat (eds.),
    <u>Perspectives</u> <u>on</u> <u>functional</u> <u>grammar</u>. Dordrecht:
    Foris. 75-88.

Duff, P. 1980. The Japanese postpositions <u>wa</u> and <u>ga</u>:
    syntactic, semantic, and pragmatic functions.
    Unpublished honours thesis. The University of
    Calgary.

Duff, P. 1984. An analysis of Mandarin and Japanese
    interlanguage. ESL 650 Term Paper. The University of
    Hawaii at Manoa.

Duff, P. 1985. A literature review and pilot study of the
    role of the Pro-drop parameter in Ll grammars and in
    second language acquisition. ESL 750 Term Paper.
    The University of Hawaii at Manoa.

Eckman, F. 1981. Markedness and degree of difficulty in
    second language learning. In <u>Proceedings</u> <u>of</u> <u>the</u> <u>fifth</u>
    <u>congress</u> <u>of</u> <u>l'Association</u> <u>internationale</u> <u>de</u>
    <u>linguistique</u> <u>apliquee.</u> Montreal, 1978. Quebec City:
    les Presses de l'Universite Laval. 115-126.

Eckman, F., L. Bell, and D. Nelson (eds.). 1984.
    <u>Universals</u> <u>of</u> <u>second</u> <u>language</u> <u>acquisition</u>. Rowley,
    Massachusetts: Newbury House.

Ferguson, C. 1975. Toward a characterization of English
    foreigner talk. <u>Anthropological</u> <u>Linguistics</u> 17,1.
    1-14.

Flynn, S. 1984. A universal in L2 acquisition based on a
    PBD typology. In F. Eckman, L. Bell, and D. Nelson
    (eds.), 75-87.

Foley, W., and R. Van Valin, Jr. 1984. Functional syntax and universal grammar. Cambridge: Cambridge University Press.

Gass, S. 1979. Language transfer and universal grammatical relations. Language Learning 29,2. 327-344.

Gass, S. 1984. A review of interlanguage syntax: language transfer and language universals. Language Learning 34,2. 115-131.

Gass, S., and J. Ard. 1984. Second language acquisition and the ontology of language universals. In W. Rutherford (ed.), 33-68.

Gass, S., and L. Selinker (eds.). 1983. Language transfer in language learning. Rowley, Massachusetts: Newbury House.

Givon, T. 1979. On understanding grammar. New York: Academic Press.

Givon, T. 1984. Universals of discourse structure and second language acquisition. In W. Rutherford (ed.), 109-136.

Gordon, D. 1985. The marking of conditionality in one learner's interlanguage: a semantic analysis. In Working Papers 4,1. Department of English as a Second Language. The University of Hawaii at Manoa. 77-102.

Greenberg, J. 1966. Some universals of language, with particular reference to the order of meaningful elements. In J. Greenberg (ed.), Universals of language. Cambridge, Massachusetts: MIT Press. 73-113.

Gruber, J. 1967, Topicalization in child language. Foundations of Language 3. 37-65. [Reprinted in D. Reibel and S. Schane (eds.), 1969. Modern studies in English. Englewood Cliffs, New Jersey: Prentice-Hall. 422-447.]

Gundel, J. 1971. Role of topic and comment in linguistic theory. Bloomington, Indiana: Indiana Linguistics Club.

Gundel, J. 1978. Stress, pronominalization and the given-new distinction. University of Hawaii Working Papers in Linguistics 10,2. 1-11.

Gundel, J. 1980. Zero NP-anaphora in Russian: a case study of topic prominence. In J. Kreiman and A. Ojeda (eds.), Papers from the parasession on pronouns and anaphora. Chicago Linguistics Society. 139-146.

Gundel, J. 1985. Universals of topic-comment structures. Paper presented at the '85 UWM Linguistics Symposium. University of Wisconsin-Milwaukee.

Gundel, J., and E. Tarone. 1983. Language transfer and the acquisition of pronominal anaphora. In S. Gass and L. Selinker (eds.), 281-296.

Haiman, J. 1978. Conditionals are topics. Language 54,3. 564-589.

Hakuta, K. 1976. A case study of a Japanese child learning English as a second language. Language Learning 26,2. 321-351.

Hatch, E. 1976. Discourse analysis, speech acts, and second language acquisition. Workpapers in TESL 10. 51-63.

Hawkins, J. 1980. On implicational and distributional universals of word order. Journal of Linguistics 16,2. 193-236.

Rilles, S. 1985. Interlanguage and the pro-drop parameter. Paper presented at the Second Language Research Forum, University of California at Los Angeles.

Hinds, J. 1978. Ellipsis and prior mention in Japanese conversation. LACUS 5. 318-335.

Hinds, J. 1983. Topic continuity in Japanese. In T. Givon (ed.), Topic continuity in discourse: a quantitative cross-language study. Amsterdam: John Benjamins. 43-93.

Howard, I., and A. Niyekawa-Howard. 1976. Passivization. In M. Shibatani (ed.), Syntax and Semantics 5. New York: Academic Press. 201-237.

Huang, J. C-T. 1984. On the distribution and reference of
empty pronouns. Linguistic Inquiry 15,4. 531-574.

Huebner, T. 1979. Order-of-acquisition vs. dynamic
paradigm: a comparison of method in interlanguage
research. TESOL Quarterly 13,1. 21-28.

Huebner, T. 1982. A longitudinal analysis of the
acquisition of English by an adult Hmong speaker.
Ph.D. Dissertation. The University of Hawaii at
Manoa.

Huebner, T. 1983. Linguistic systems and linguistic
change in an interlanguage. Studies in Second
Language Acquisition 6,1. 33-53.

Huebner, T. 1985. System and variability in interlanguage
syntax. Language Learning 35,2. 141-163.

Hunt, K. 1970. Syntactic maturity in schoolchildren and
adults. Monographs of the Society for Research in
Child Development 35,1.

Hunt, K. 1977. Early blooming and late blooming syntactic
structures. In C. Cooper and L. Odell (eds.),
Evaluating writing: describing, measuring, judging.
Urbana: National Council of Teachers of English.

Hwang, M. 1985. Topic continuity/discontinuity in Korean
narrative. Korean Linguistics 10,1.

Jordens, P. 1983. Discourse functions in interlanguage
morphology. In S. Gass and L. Selinker (eds.), 327-
357.

Jordens, P., and E. Kellerman. 1981. Investigations into
the "transfer strategy" in second language learning.
In Proceedings of the 5th congress of l'Association
internationale de linguistique appliquee. Montreal
1978. 195-215.

Keenan, E.L. 1976. Towards a universal definition of
"subject." In C. Li (ed.), 303-334.

Kellerman, E. 1983. Now you see it, now you don't.
In S. Gass and L. Selinker (eds.), 112-134.

Kitagawa, C. 1982. Topic constructions in Japanese.
Lingua 57. 175-214.

Kleinmann, H. 1977. Avoidance behavior in adult second language acquisition. Language Learning 27.1. 93–107. [Reprinted in R. Robinett and J. Schachter (eds.), second language learning, contrastive analysis, error analysis, and related aspects. Ann Arbor: The University of Michigan Press. 363–377.]

Kodama, T. 1981. Topic and scene–setter in English and Japanese: a contrastive analysis based on daughter-dependency grammar. In T. Konishi (ed.), Studies in grammar and language. Tokyo: Kenkyusha.

Kress, G. (ed.). 1976. Halliday: system and function of language. London: Oxford University Press.

Kuno, S. 1973. The structure of the Japanese language. Cambridge: MIT Press.

Larsen–Freeman, D. 1983. Assessing global second language proficiency. In H. Seliger and M. Long (eds.), 284–304.

Lehmann, W. 1973. A structural principle of language and its implications. Language 49,1. 47–66.

Li, C. (ed.). 1976. Subject and topic. New York: Academic Press.

Li, C., and S. Thompson. 1976. Subject and topic: a new typology of language. In C. Li (ed.), 457–490.

Li, C., and S. Thompson. 1978. An exploration of Mandarin Chinese. In W. Lehmann (ed.), Syntactic typology. Austin: University of Texas Press. 223–266.

Li, C., and S. Thompson. 1979. Chinese pronouns and zero–anaphora in Chinese discourse. In T. Givon (ed.), Syntax and Semantics 12. New York: Academic Press. 311–335.

Li, C., and S. Thompson. 1981. Mandarin Chinese: a functional reference grammar. Berkeley: university of California Press.

Lujan, M., L. Minaya, and D. Sankoff. 1984. The universal consistency hypothesis and the prediction of word order acquisition stages in the speech of bilingual children. Language 60,2. 343–371.

191

Magretta, W. 1977. Topic-comment structure and linguistic
    theory: A functional approach. Ph.D. Dissertation.
    The University of Michigan.

Mallinson, G., and B. Blake. 1981. Language typology.
    Amsterdam: North-Holland Publishing Company.

Martin, S. 1975. A reference grammar of Japanese.
    New Haven: Yale University Press.

Meisel, J. 1980. Linguistic simplification. In S. Felix
    (ed.), Second language development: trends and
    issues. Tubingen: Narr. 13-40.

Reinhart, T. 1982. Pragmatics and linguistics: an
    analysis of sentence topics. Bloomington, Indiana:
    Indiana University Linguistics Club.

Ross, J. 1967. Constraints on variables in syntax.
    Ph.D. Dissertation. MIT.

Rutherford, W. 1977. Modern English. Second Edition?
    Volume 2. New York: Harcourt Brace Jovanovich.

Rutherford, W. 1982. Markedness in second language
    acquisition. Language Learning 32,1. 85-108.

Rutherford, W. 1983. Language typology and language
    transfer. In S. Gass and L. Selinker (eds.), 358-
    370.

Rutherford, W. (ed.). 1984a. Language universals and
    second language acauisition. Amsterdam:
    John Benjamins Publishing Company.

Rutherford, W. 1984b. Description and explanation in
    interlanguage syntax: state of the art. Language
    Learning 34,3. 127-153.

Rutherford, W., and R. Altman. 1985. Discourse competence
    in L2 acquisition: word order. Unpublished ms. The
    University of Southern California.

Sasaki, Y. 1985. Two new waves of language acquisition
    research: the Pienemann-Johnston model and the Bates-
    MacWhinney model. Linguistics 640 Term Paper. The
    University of Hawaii at Manoa.

Schachter, J. 1974. An error in error analysis. <u>Language Learning</u> 24,2. 205-214.

Schachter, J. 1983. A new account of language transfer. In S. Gass and L. Selinker (eds.), 98-111.

Schachter, J., and B. Hart. 1979. An analysis of learners' production of English structures. <u>Georgetown University Papers on Languages and Linguistics</u> 15. 18-75.

Schachter, J., and W. Rutherford. 1979. Discourse function and language transfer. <u>Working Papers on Bilingualism</u> 19. 1-12.

Schmidt, M. 1980. Coordinate structures and language universals in interlanguage. <u>Language Learning</u> 30,2. 397-416.

Schumann, J. 1979. The acquisition of English negation by speakers of Spanish: a review of the literature. In R. Andersen (ed.), <u>The acquisition and use of Spanish and English as first and second languages</u>. Washington, D.C.: TESOL. 3-32.

Schumann, J. Forthcoming. Utterance structure in basilang speech. To appear in G. Gilbert (ed.), <u>Pidgin and creole languages: essays in memory of John E. Reinecke</u>. University of Hawaii Press.

Seliger, H. 1984. Processing universals in second language acquisition. In F. Eckman, L. Bell, and D. Nelson (eds.), 36-50.

Seliger, H., and M. Long. 1983. <u>Classroom oriented research</u>. Rowley, Massachusetts: Newbury House.

Slobin, D. 1977. Language change in childhood and in history. In J. Macnamara (ed.), <u>Language learning and thought</u>. New York: Academic Press.

Smith, D. 1978. Mirror images in Japanese and English. <u>Language</u> 54,1. 78-122.

Swain, M., N. Naiman, and G. Dumas. 1972. Aspects of the learning of French by English-speaking 5 year olds. Paper presented at AILA, Copenhagen.

Tai, J. H-Y. 1982. Relevant categorical distinctions in Chinese. In K. Tuite, R. Schneider, and R. Chametzky (eds.), Papers from the eighteenth regional meeting. University of Chicago: Chicago Linguistic Society. 495-506.

Thompson, S. 1978. Modern English from a typological point of view: some implications of the function of word order. Linguistische Berichte 54,78. 19-35.

van Dijk, T. 1981. Studies in the pragmatics of discourse. The Hague: Mouton.

White, L. 1984. Implications of parametric variation for second language acquisition: an investigation of the pro-drop parameter. Revised version of paper presented at the Colchester Second Language Acquisition Workshop on Experimental Approaches to Second Language Acquisition. University of Essex. January, 1984.

White, L. 1985. The Pro-drop parameter in adult second language acquisition. Language Learning 35,1. 47-62.

Xu, L., and D. Langendoen. 1985. Topic structures in Chinese. Language 61,1. 1-27.

Zobl, H. 1982. A direction for contrastive analysis: The comparative study of developmental sequences. TESOL Quarterly 16,2. 169-183.

Zobl, H. 1983. L1 acquisition, age of L2 acquisition, and the learning of word order. In S. Gass and L. Selinker (eds.), 205-221.

Zobl, H. 1984. Uniformity and source language variation across developmental criteria. In W. Rutherford (ed.), 185-218.

Zobl, H. 1985a. Typological control in interlanguage. Unpublished ms. The University of Hawaii at Manoa and The University of Moncton.

Zobl, H. 1985b. A functional approach to the attainability of typological targets in L2 acquisition. Unpublished ms. The University of Hawaii at Manoa and The University of Moncton.