

**PLACEMENT TESTING AND MORPHOSYNTACTIC DEVELOPMENT
IN SECOND LANGUAGE LEARNERS OF ENGLISH**

by

Patti A. Spinner

B.A., Rutgers University, 1995

M.A., Ohio State University, 1999

Submitted to the Graduate Faculty of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Patti A. Spinner

It was defended on

July 23, 2007

and approved by

Dr. G. Richard Tucker, Paul Mellon University Professor, Carnegie Mellon University

Dr. Pascual Masullo, Associate Professor, University of Pittsburgh

Dr. Yasuhiro Shirai, Professor, University of Pittsburgh

Dissertation Advisor: Dr. Alan Juffs, Associate Professor, University of Pittsburgh

Copyright © by Patti Spinner

2007

**PLACEMENT TESTING AND MORPHOSYNTACTIC DEVELOPMENT
IN SECOND LANGUAGE LEARNERS OF ENGLISH**

Patti Spinner, M.A.

University of Pittsburgh, 2007

ABSTRACT: The primary purpose of this dissertation is to discover whether two current proposals for specific indicators of morphosyntactic development can successfully predict the placement of second language learners of English (ESL learners) in an intensive English program. This research is important because most of the placement/proficiency tests that are currently in use do not include a clear, empirically-tested theory of how second language learners (L2 learners) acquire the morphosyntax of the target language, which is one essential component of L2 proficiency. In order to determine which morphosyntactic elements could be included in a new assessment measure, I examined semi-spontaneous oral production data from 48 ESL learners of mixed L1 background at an intensive English program at the University of Pittsburgh. The measures examined and methodology used were based primarily on Young-Scholten, Ijuin, & Vainikka's (2005) Organic Grammar and Pienemann's (2003) Rapid Profile, two proposals that intend to account for L2 learner development. In order to test the proposals of each, I created implicational tables based on the production data. It was found that Organic Grammar could not fully account for the order of emergence of morphosyntactic features in these data. While Rapid Profile made more accurate predictions, the predictions were not useful in distinguishing between learners at intermediate and advanced levels. Despite these problems, it was possible to combine the results from the Organic Grammar and Rapid Profile tables to produce a new table describing the order of emergence of morphosyntactic forms. It is possible

that this table can be integrated into current scale measures of placement/proficiency, such as the ACTFL scale. A preliminary proposal for such a combined measure is proposed; however, further empirical research is necessary in order to determine the effectiveness and accuracy of the scale.

TABLE OF CONTENTS

PREFACE.....	XIV
1.0 INTRODUCTION.....	1
2.0 LITERATURE REVIEW.....	5
2.1 PROFICIENCY/PLACEMENT TESTING.....	5
2.1.1 The nature of proficiency	5
2.1.2 Widely used placement and proficiency tests.....	16
2.1.2.1 Large-scale proficiency measures	16
2.1.2.2 Other proficiency and placement measures.....	20
2.1.2.3 The use of rubrics	23
2.1.3 An alternative to traditional tests: Developmental approaches	27
2.2 ORGANIC GRAMMAR.....	30
2.2.1 The basis of Organic Grammar: Generative syntax	30
2.2.1.1 Lexical categories and functional categories.....	30
2.2.1.2 Functional categories: IP and CP.....	32
2.2.2 The origin of Organic Grammar: L1 acquisition research.....	37
2.2.3 Minimal Trees Theory	39
2.2.3.1 The L2 Initial State: A transferred VP.....	39
2.2.3.2 Lack of functional categories at the initial state	43

2.2.3.3	Gradual development of phrase structure	48
2.2.4	A placement test based on Minimal Trees.....	53
2.3	PROCESSIBILITY THEORY AND RAPID PROFILE.....	55
2.3.1	The theoretical basis of Processability Theory: LFG	55
2.3.2	The origins of Rapid Profile: Processability Theory	58
2.3.3	A processing-based proficiency measure	66
2.3.4	The Multidimensional model	70
2.3.5	Summary of section 2.....	74
3.0	THE STUDY.....	76
3.1	METHODOLOGY	76
3.1.1	The environment	76
3.1.2	The data	79
3.1.3	Participants.....	82
3.1.4	Language background	85
3.1.4.1	Arabic.....	86
3.1.4.2	Korean	92
3.1.4.3	Chinese (Mandarin).....	98
3.1.4.4	Spanish.....	104
3.1.4.5	Summary of language descriptions	111
3.1.5	Measures	112
3.1.5.1	Examining Organic Grammar: Features of VP, IP and CP.....	114
3.1.5.2	Testing Pienemann’s Rapid Profile.....	136
3.1.5.3	Statistical procedures.	139

3.1.5.4	Interviews	144
3.1.6	Research questions.....	150
3.2	RESULTS	154
3.2.1	Testing Organic Grammar.....	154
3.2.2	Testing Rapid Profile.....	167
3.2.3	Other measures	172
3.2.4	Interview results.....	174
3.2.4.1	Overall development.....	174
3.2.4.2	Error measures	175
3.2.4.3	Complexity measures.....	180
3.2.5	L1 influence	182
4.0	DISCUSSION	185
4.1	ORGANIC GRAMMAR AND THE PATH OF DEVELOPMENT	185
4.1.1	Emergence of morphosyntactic features.....	185
4.1.2	Morphological inflection and developmental order.....	190
4.2	RAPID PROFILE AND THE PATH OF DEVELOPMENT	193
4.3	A DIFFERENT PATH OF DEVELOPMENT	196
4.4	FIRST LANGUAGE INFLUENCE	200
4.5	INFLUENCE OF MOTIVATION AND EXPOSURE TO ENGLISH	205
4.6	“COMPLEXITY” IN THE PATH OF DEVELOPMENT	207
5.0	CONCLUSION	214
5.1	DEVELOPMENTAL FEATURES AND PLACEMENT TESTS	214
5.2	DIRECTIONS OF FUTURE RESEARCH	221

5.3	RECOMMENDATIONS FOR INSTRUCTION AND ASSESSMENT.....	223
APPENDIX A	227
APPENDIX B	232
REFERENCES	235

LIST OF TABLES

Table 1. MTELP grammar and vocabulary questions (Michigan, 2007)	21
Table 2. The German subject-verb agreement paradigm for the verb <i>schwimmen</i> , ‘to swim’	44
Table 3. Organic Grammar (Young-Scholten, Ijuin, & Vainikka 2006, p. 9)	54
Table 4. Rapid Profile stages (Pienemann, 2003).....	68
Table 5. New student level placement	78
Table 6. Participants with L1 and entry level	83
Table 7. Biographical information of participants.....	84
Table 8. Abbreviations used in table.....	155
Table 9. Implicational Table: Predictions of Organic Grammar	156
Table 10. Coefficients of reproducibility for elements in Organic Grammar.....	158
Table 11. OG implicational table revised for greatest scalability and reproducibility	160
Table 12. Coefficients of reproducibility for elements in revised OG table.....	162
Table 13. Implicational Table revised with cutoff points of 80% for inflection	164
Table 14. Revised OG table with morphological measures removed.....	166
Table 15. Abbreviations and symbols used in Rapid Profile table	168
Table 16. Implicational table: Predictions of Rapid Profile	169
Table 17. Coefficients of reproducibility for elements in the Rapid Profile table.....	171

Table 18. Sample size and vocabulary measures for learners	173
Table 19. Morphosyntactic features of integrative and less integrative learners.....	174
Table 20. Copula suppliance of integrative and less integrative learners.....	175
Table 21. Subject suppliance for integrative and less integrative learners	177
Table 22. Determiner suppliance in integrative and less integrative learners	177
Table 23. Errors per clause in highly integrative and less integrative speakers	179
Table 24. Words per AS-unit for integrative and less integrative speakers.....	180
Table 25. Clauses per AS-unit for integrative and less integrative learners	181
Table 26. SOC of subjects for learners who omitted subjects	183
Table 27. Learners' use of relative clauses	184
Table 28. Organic Grammar (Young-Scholten, Ijuin, & Vainikka 2006, p. 9).....	186
Table 29. Implicational table combining measures of RP and OG for maximum scalability ...	198
Table 30. Learners' use of subjects.....	202
Table 31. Learners' use of relative clauses	204
Table 32. Learners' use of modals in spontaneous production data.....	211
Table 33. Learners' use of adverb clauses in spontaneous production data	212
Table 34. Selected descriptions from ACTFL scale (1999)	217
Table 35. Morphosyntactic elements corresponding to ACTFL descriptions	219

LIST OF FIGURES

Figure 1. Bachman's model of language competence (1990, p. 87).....	10
Figure 2. Bachman's model of communicative competence (1990, p. 85)	11
Figure 3. A sentence with a modal.....	32
Figure 4. Subject in SPEC of IP.....	33
Figure 5. Sentence with CP complement	34
Figure 6. A yes/no question	35
Figure 7. A wh-question	36
Figure 8. Movement in German matrix clauses.....	41
Figure 9. A learner utterance at the VP stage	49
Figure 10. A learner utterance at the FP stage	50
Figure 11. A learner utterance at the AgrP stage.....	51
Figure 12. A learner utterance at CP (final) stage	52
Figure 13. An example of c-structure in LFG.....	56
Figure 14. An ideal implicational table.....	140
Figure 15. An implicational table with two errors	140
Figure 16. Measures used to explore the Multidimensional Model.....	150
Figure 17. Copula suppliance of integrative and less integrative learners	176

Figure 18. Determiner suppliance by integrative and less integrative learners	178
Figure 19. Errors per clause in highly integrative and less integrative speakers	179
Figure 20. Words per AS-unit for integrative and less integrative learners	181
Figure 21. Clauses per AS-unit for integrative and less integrative learners.....	182

PREFACE

This project would not have been possible without the support of many people. Let me begin by extending many thanks to my fellow graduate students, especially Maeve Eberhardt, Guillermo Rodriguez, Brian Brubaker and Wendy Martelle, whose support and collegiality helped me to make it to this point. As for the current project, first I would like to acknowledge the National Science Foundation, from whom I received funding via award number SBE-0354420 to the Pittsburgh Science of Learning Center. At various stages of this project, I have also been assisted by the expertise and advice of Lois Wilson, Betsy Davis, Dawn McCormick, and M. Christine O'Neill, as well as teachers in the English Language Institute, especially Chris Ortiz and Tamar Bernfeld. Ben Friedline was invaluable in transcribing interviews and helping to analyze the data. I would also like to extend thanks to Shanwen Yu and Lashanda Lemmon, who helped me obtain my data and kept good humor despite a few last-minute requests for assistance. Various people have generously offered their time and talents to review sections of this dissertation: Ethan Pullman, Ebru Turker, and Guillermo Rodriguez. I'd especially like to thank Greg Mizera, who kindly offered to help me through the final stages of this project and who suggested new perspectives on the results. Sincere thanks to the members of my committee, Pascual Masullo, Dick Tucker, and Yasuhiro Shirai, for sharing their expertise and for helping me take this project in new directions. Finally, I'd like to thank Alan Juffs, who inspired me to enter this field, and who has both supported my career and challenged my intellect for the past

six years. His thoughtful comments on a number of drafts of this dissertation were immeasurably helpful.

I would like to dedicate this dissertation to my mother, who always wanted me to succeed, and who would have been very proud.

1.0 INTRODUCTION

The primary purpose of this dissertation is to discover whether two current proposals for specific indicators of morphosyntactic development can successfully predict the placement of second language learners of English (ESL learners) in an intensive English program. This research is important because most of the placement/proficiency tests that are currently in use do not include a clear, empirically-tested theory of how second language learners (L2 learners) acquire the morphosyntax of the target language, which is one essential component of L2 proficiency. Generally, these proficiency measures either test knowledge of a large number of grammatical forms in a multiple-choice format, or decline to refer to specific grammatical forms at all. It is desirable to create placement/proficiency tests which are based in sound theories of L2 development.

Recently, several researchers have proposed L2 assessment tests which are based in theories of the development of morphosyntax. Young-Scholten, Ijuin, & Vainikka (2005) propose Organic Grammar, a placement test based on Vainikka & Young-Scholten's (1994; 1996a,b; 1998) Minimal Trees theory. Minimal Trees theory, which has its basis in Generative syntactic theory, claims that L2 learners have only a Verb Phrase (VP) in the initial state of acquisition and must gradually acquire functional projections (and the forms and structures which are associated with them). Another L2 assessment tool is Pienemann's (2003) Rapid Profile, which is an account of the L2 development of morphosyntax based on Pienemann's

(1998) Processability Theory. According to Rapid Profile, the relative difficulty of a morphosyntactic element depends on the distance over which grammatical information must be related: within a phrase, between phrases, or at the beginning or ends of clauses.

Any placement test that is based strictly on morphosyntactic development is assuming a kind of “unitary trait hypothesis” (c.f., Oller, 1976). In other words, the assumption is that morphosyntactic development underlies all other types of linguistic behavior—or at the least, the types of linguistic behavior that are relevant to language teaching. It is not yet clear to what extent this assumption is correct.

Placement/proficiency tests which are based strictly on morphosyntactic development are on one end of the spectrum; on the other end are tests which make little reference to specific grammatical forms at all. In the U.S., the most widely-used measure of this type is the ACTFL (American Council of Teachers of Foreign Languages) proficiency scale. The ACTFL scale describes the capabilities of L2 learners from the beginning stages and as they become more adept at communicating. For example, some of the descriptors refer to the types of social situations the learner can participate in, the types of topics the learner can discuss, and the level of vocabulary the learner can produce. The rubric also refers to the number of errors a learner makes and how easily she can be understood by native speakers. The descriptors are unfortunately fairly vague and relative; the only specific grammatical forms that are referred to in the scale are tense and aspect marking. The score a learner receives on an exam based on the ACTFL scale is therefore quite subjective, based largely on the assessor’s individual interpretation of the descriptors.

The goal of this dissertation is to begin to develop a measure which is a middle ground between these two extremes: that is, a proficiency test that makes use of empirically tested,

objective morphosyntactic development measures, but that also allows for the inclusion of the types of proficiency indicators that are used in scales such as the ACTFL rubric. It is hoped that such a measure would retain the “best of both worlds,” while rejecting aspects of the assessment tools that are oversimplified, inaccurate, or ambiguous.

The targeted learner group for this study is adult ESL learners in intensive English programs (IEPs) in the United States or other English-speaking countries. These are learners over the age of 17 from various L1 backgrounds who have generally had several years of English schooling before entering the program. They are not beginners, but they are not yet at a proficiency level which allows them to accomplish their goals (entering an English-speaking university, using English for business, etc.). This group was selected for the study because of the large number of IEPs in English-speaking countries and their great need for accurate and reliable proficiency tests.

The first step in developing such a placement/proficiency measure is to examine the orders of emergence of morphosyntactic forms that are predicted by Organic Grammar and Rapid Profile. Both of these proposals are still controversial, and it remains to be seen if they can account for the development of learners of various L1 backgrounds in an intensive English program. If a clear path of morphosyntactic development can be identified, then it should be able to be used as part of a placement/proficiency measure.

It was discovered that, although neither the predictions of Organic Grammar nor those of Rapid Profile were accurate and fine-grained enough to be used as placement tests on their own, the morphosyntactic elements that they measure are acquired in a predictable way. It is possible that these elements can be integrated into current scale measures of placement/proficiency, such as the ACTFL scale. A preliminary proposal for such a combined measure is proposed;

however, further empirical research is necessary in order to determine the effectiveness and accuracy of the scale.

This dissertation is organized as follows. First, section 2.0 presents background information on proficiency/placement testing (section 2.1), Organic Grammar (section 2.2) and Rapid Profile (section 2.3). Section 3.0 presents the current study, including the methodology used (section 3.1), and the results (section 3.2). Section 4 is a discussion of the results, and Section 5 provides the proposal for the new placement/proficiency rubric, as well as directions for further research.

2.0 LITERATURE REVIEW

2.1 PROFICIENCY/PLACEMENT TESTING

2.1.1 The nature of proficiency

The goal of placement tests in language programs is to divide students into (relatively) homogenous groups for the purpose of class assignment. Language classes can be organized according to students' interests, first languages, or linguistic or academic goals, but the most typical criterion for dividing students into classes is linguistic ability in the target language. For this reason, placement tests generally measure second language proficiency.

Accurate and reliable placement tests are crucial to the successful functioning of an English language program. Students who are placed into classes that are too advanced may not be able to follow classroom discourse or participate in class activities, and therefore may fail to improve their language ability. There is a cost to instructors, as well, who may need to spend extra time trying to help these students. Conversely, students who are placed into classes that are too basic will waste time covering familiar material and may become frustrated and bored.

In order to place students into groups by proficiency, it is first necessary to determine the nature of proficiency itself. Unfortunately, this task has proven to be surprisingly difficult, and researchers are still divided. Some have claimed that proficiency is essentially undefinable; for

instance, both Vollmer (1981) and Pienemann (1985) have written that proficiency can only be defined as that which is measured by proficiency tests. While this view is certainly defensible, it unfortunately does not help to provide a foundation for the construction of a proficiency test. Testing specialists have therefore been seeking a clear account of proficiency that can serve as a theoretical guide in the construction of assessment measures. Over the course of more than thirty years, there has been steady progress in this regard.

In the early 1980s, many researchers thought that it was possible to measure proficiency as a single, indivisible factor that underlies all types of language performance. This belief was based largely on work on the “Unitary Trait Hypothesis” by Oller (1976), who conducted factor-analysis studies on ESL students’ performance on language tests and claimed to find evidence for the existence of an underlying linguistic ability that determined students’ performance in all content and skill areas: the so-called “g-factor”. The g-factor, which was supposedly related to overall intelligence, was defined as a learner’s ability to integrate grammatical, pragmatic, contextual, and lexical skills. Oller claimed that cloze tasks were ideal for measuring the g-factor, and could serve to provide a general proficiency score without the need for a battery of separate measures.

Oller’s proposals eventually met with fierce criticism, primarily revolving around the statistical procedures he had used. For instance, Bachman & Palmer (1981, 1982) used a new method (the Multitrait-Multimethod Matrix) to show that speaking and reading proficiencies were independent. Similarly, Vollmer & Sang (1983) pointed out errors in Oller’s methods and reanalyzed his data; they found that multifactor solutions fit the data more precisely. Oller himself withdrew his claim in 1983, saying that the Unitary Trait Hypothesis—at least in its strong form—was wrong.

In most models that followed, proficiency was considered to consist of a number of separate, if related, skills (e.g., Bachman & Palmer, 1981; Upshur & Homburg, 1983). For instance, a learner's ability to perform on speaking tasks may be better or worse than his ability to perform on writing or reading tasks. However, many questions remain about these skills. For instance, it is not yet clear how many components there are or to what extent they are interrelated. Similarly, it is also not known whether there is a single "set" of language skills that are used in differing amounts in different contexts of language use, or whether each context requires a unique set of skills (Bachman & Palmer, 1981).

Oller's model receded in prominence as a new mindset began to dominate thought about how language is learned and used: the communicative approach. Researchers began to focus on language as a tool for communication, rooted in social interaction. As Berns (1984, p. 5) expresses it:

Language is interaction; it is interpersonal activity and has a clear relationship with society. In this light, language study has to look at the use (function) of language in context, both its linguistic context (what is uttered before and after a given piece of discourse) and its social, or situational context (who is speaking, what their social roles are, why they have come together to talk).

In the 1980s, researchers started to include the communicative dimension of language in models of language proficiency. Canale & Swain (1980) built on work by earlier researchers (e.g., Habermas, 1970; Hymes, 1971), to produce their influential formulation of communicative competence. For them, language proficiency was not simply the linguistic ability of a learner, but also—crucially—the way the learner employs linguistic knowledge when communicating in a variety of contexts. Canale & Swain's (1980) model includes four areas that are encompassed

in proficiency: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Grammatical competence is knowledge and mastery of the linguistic code, including vocabulary use, pronunciation, spelling, morphological rules and syntactic rules. A learner with good grammatical competence will have high linguistic accuracy on word forms. The second aspect of proficiency, sociolinguistic competence, is the use of appropriate language in a variety of social contexts, i.e., the ability to behave in a linguistically appropriate manner so as to achieve the desired results with the person or people being addressed. It includes the ability to select the correct forms for formal versus informal contexts and the ability to choose forms that are considered polite by native speakers. For instance, if a learner wants to ask someone to open a window, she needs to know not only how to formulate a grammatical utterance, but also how to phrase the request so as to avoid offending the listener. The third aspect of proficiency, discourse competence, is the ability to choose meanings and forms that together create a unified and organized text or message. For instance, in order for a speaker to be able to express herself meaningfully in a phone conversation, she needs to master the use of greetings, introduction to a topic, and closure. Also included in discourse competence is the ability to make a persuasive argument with the use of cohesion devices and argumentation, whether in written or oral form. The last aspect of proficiency is strategic competence, the ability to compensate for lack of second language knowledge, such as with the use of circumlocution and body language.

Another prominent model of communicative competence, presented in Bachman (1990), is similar to Canale & Swain's, but has a slightly different organization. Bachman breaks communicative competence into three parts: language competence, strategic competence, and psychophysiological mechanisms. Language competence is described the most completely, and includes any type of expression that requires a knowledge of linguistic forms: grammatical

ability, skills in cohesion and organization (textual competence), and pragmatic and sociolinguistic skills (illocutionary competence and sociolinguistic competence). Each of these areas is finely articulated by Bachman. Grammatical ability includes use of vocabulary, morphological forms, syntactic structure, and pronunciation or spelling. Textual competence includes linguistic organization and the ability to formulate an argument. Illocutionary competence is divided into four separate areas: manipulative, heuristic, imaginative, and ideational. The ability to perform manipulative functions is the ability to influence the world around the speaker. The heuristic component refers to the ability to learn from language, such as by reading or memorizing information. The imaginative function of language gives language users the ability to bring humor or aesthetic pleasure into their lives through language. Finally, the learner must be able to use language to exchange information in the world, using the ideational function of language. The final component of language competence in Bachman's model is sociolinguistic competence. Sociolinguistic competence refers to a learner's ability to perceive and understand various dialects or varieties of a language, to use and understand appropriate forms for both casual and formal situations, to use natural-sounding expressions, and to understand cultural references.

A diagram of Bachman's model of language competence is provided in Figure 1.

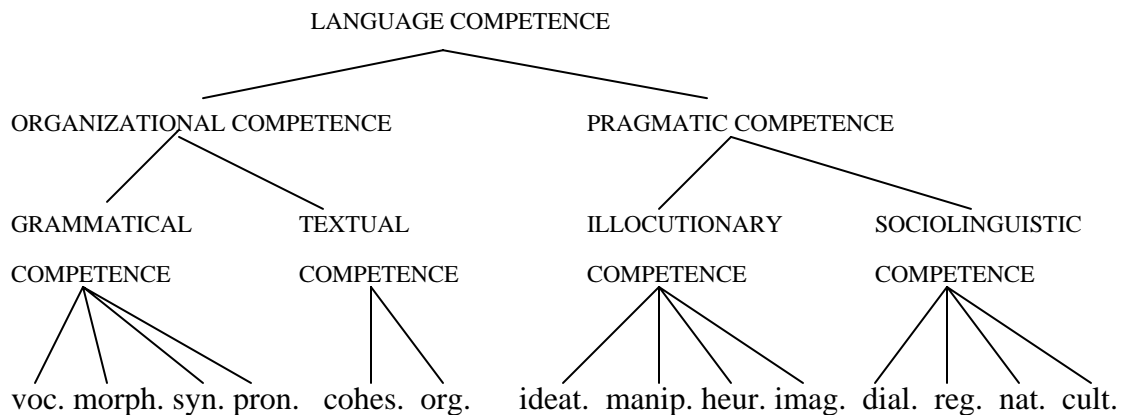


Figure 1. Bachman's model of language competence (1990, p. 87)

A second aspect of communicative competence in this model is strategic competence, which includes learners' ability to use rhetorical devices to enhance the effect of their speech or writing, and also their ability to compensate for lack of second language knowledge with the use of circumlocution and body language. As a final aspect of communicative competence, Bachman includes psychophysiological mechanisms—the neurological and physiological processes that occur when producing and comprehending language—as a part of communicative language use, because they are required for communication to take place. Figure 2 below shows Bachman's model of communicative competence.

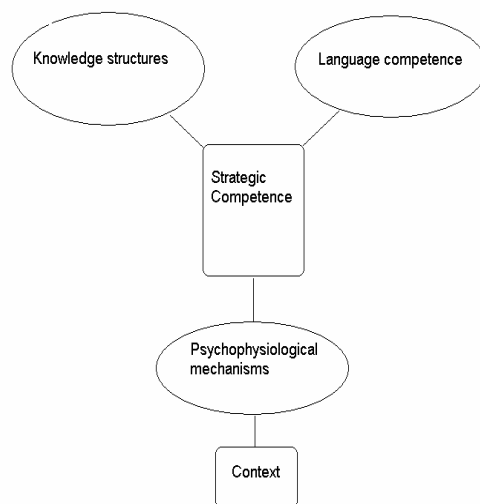


Figure 2. Bachman's model of communicative competence (1990, p. 85)

While the basic concept of communicative competence has remained unchanged since Canale & Swain and Bachman's models, there have been some refinements. For instance, Savignon (2001, 2002) argues that sociolinguistic competence extends beyond the use of appropriate linguistic forms. To communicate successfully, learners require an understanding of the social and cultural contexts in which communication takes place, including the roles of the speakers, the knowledge that they share or don't share, and the purposes of the exchange. Savignon proposes the term *sociocultural competence* to replace sociolinguistic competence in a model of communicative competence.

Additionally, researchers are becoming more interested in the use of language strategies, that is, techniques or behaviors that learners use to facilitate communication in a second language and to continually improve their second language functioning. For example, a learner with good reading strategies may skim the title and subheadings of an article, so as to have an idea about the content of the article before beginning to read. A number of studies have examined the relationship between learners' use of strategies and their language proficiency, and

it may be worthwhile to include the use of learning strategies in a model of proficiency (e.g., Oxford, 1990; Nisbet, Tindall, & Arroyo, 2005).

In the work by these and other researchers, the concept of language proficiency as communicative competence has been quite well defined. However, problems arise when trying to measure it. Taylor (1988) points out that it is impossible to directly measure competence, since only performance can be observed. Spolsky (1989) agrees, saying that even though researchers generally refer to learners' second language competence, tests are more likely to measure language use. Bachman & Savignon (1986) and Taylor (1988) advocate replacing the term *competence* with *communicative language proficiency* and *communicative proficiency*, respectively, to refer to that which can be measured by tests. In this light, proficiency is thought of as the ability to make use of competence.

Once we begin to think about measuring language performance, certain problems arise. For one thing, it has been shown that learners vary greatly in their level of accuracy; that is, the same learner may appear to be highly proficient in one situation, but quite inaccurate in another. Tarone (1983) claims that the crucial variable is the amount of attention that learners pay to the form of their utterance. Her explanation is that learners have a range of styles which range between the extreme "vernacular" style or extreme "careful" style. Vernacular style is used in unattended speech (such as during relaxed socializing), while the careful style might be elicited through a grammaticality judgment task. To give an example, Dickerson (1975) found that Japanese-speaking learners produced much more accurate English pronunciation in classroom activities than they did in social contexts outside of the classroom.

Complicating the matter even further, individual learners may tend towards a vernacular or careful style in all speech activities. For instance, Foster & Skehan (1994) found that some

learners generally have a more accuracy-oriented approach to language use, while others have a more fluency-oriented approach. Similarly, Meisel, Clahsen & Pienemann (1981) and Clahsen, Meisel & Pienemann (1983) argue that learners who are interested in “integrating” themselves into the target culture are more likely to focus on producing correct forms than those who prefer to keep themselves separated from the target culture.¹ Another possibility is that learners from certain cultures may be more likely to use an accuracy-oriented approach than others.

The question for test designers is: which style, vernacular or careful, represents a learner’s proficiency, and which is more important? The question has no clear answer, so test designers must make difficult decisions about the goals of the test, and the goals of the learner. In fact, according to one view, proficiency is meaningless without taking into account the particular needs of a learner or group of learners, since arguably every learner needs English for a specific purpose (Jarvis, 1986). For instance, a learner of French who plans to spend a week as a tourist in Paris needs a very different type of proficiency than a learner of French who plans to study at the Sorbonne.

Another major problem for test developers is knowing how to design the test questions or elicit language performance. Any measuring tool will be indirect, that is, it will not directly test a speaker’s competence or even overall performance. As Clark (1978) puts it,

The most direct procedure for determining an individual’s proficiency in a given language would simply be to follow that individual surreptitiously over an extended period of time, observing and judging the adequacy of the performance in the language-use areas in question...It is clearly impossible, or at least highly impractical, to administer a test of this type in the usual language learning

¹ This argument is part of the Multidimensional Model, which will be discussed in Section 2.3.

situation. Nonetheless, the development of proficiency measurement procedures that can properly be considered 'direct' must be based on approximating, to the greatest possible extent within the necessary constraints of testing time and facilities, the specific situations in which the proficiency is called upon in real life (p. 23).

In order to approximate real-life use, tests would be interaction-based; they would be unpredictable, as conversation and life can be; they would take place in a real communicative context (i.e., not a testing environment); they would use authentic language (not simplified), and serve a real communicative purpose, among other things (Alderson, 1981).

Testing specialists generally agree that proficiency and placement measures need to have certain qualities to be effective. The following qualities are based on Bachman (1990):

1. The test should be *reliable*; i.e., a student should receive a similar score on the test every time she takes it.
2. The test should be *valid*; i.e., the test should measure communicative competence and nothing else.
3. The test should be *authentic*; i.e., language should approximate native speech.
4. The test should be as *direct* as possible; i.e., it should attempt to approximate the skill it is measuring as closely as possible.
5. The test should be *complete*; i.e., it should cover all aspects of language ability that are important for a particular goal.
6. The test should be as *precise* as possible; i.e., there should be clear answers to all questions and rubric descriptors should be specific and detailed.

7. According to many who support communicative language testing, the test should have *interactive* elements; i.e., it should involve the possibility for give-and-take meaning interaction (Stoynoff & Chapelle, 2005).

While it may not be possible to achieve all of these goals, it is certainly the case that certain types of measures come closer than others. For example, discrete-point testing has been strongly criticized for failing to properly represent learners' competence. As early as 1961, John Carroll argued that discrete-point tests fail because it is not possible to measure proficiency in discrete units. Rather, Carroll advocated integrative tests in which learners have to use multiple language skills at once, claiming that these tests are more natural and therefore better represent learners' true proficiency. Obviously, language tests that require learners to supply grammar rules or perform complex language tasks (such as changing passive sentences to active sentences, or forming questions from statements) also lack "naturalness." However, multiple-choice tests are objective, easy to score, and cost-effective, so they may be the only viable option in some circumstances.

Those who must construct language tests are faced with these and other difficult challenges. Test designers cannot wait for the issues to be resolved; placement and proficiency tests are needed now. In the next section, various placement and proficiency tests will be discussed with the goal of learning how close they come to measuring communicative competence.

2.1.2 Widely used placement and proficiency tests

In order to satisfy the considerable need for placement and proficiency measurements in thousands of (intensive) English programs around the world, a large number of institutions have produced and marketed tests of various length, methodology, purpose, and quality. In the following sections, the most commonly used placement and proficiency tests are briefly reviewed in respect to their usefulness as a placement test for intensive English programs.

2.1.2.1 Large-scale proficiency measures

The Educational Testing Service (ETS) offers two major proficiency tests: the recently redesigned TOEFL-iBT (The Test of English as a Foreign Language, Internet based) and the TOEIC (Test of English in International Communication). An older test, the TSE (Test of Spoken English), is gradually being phased out and will be replaced by speaking component of the TOEFL-iBT.

The TOEFL-iBT claims to be “a measure of communicative language proficiency in English, and focuses on academic language and the language of university life” (Chapelle, Grabe & Burns, 1997). Generally speaking, the TOEFL is not advertised as a placement test (although, as a measure of proficiency, it is sometimes used as one); rather it is a test of English for academic purposes, and is usually taken by learners who wish to begin study at an English-speaking university. A score of roughly 80 on the newly redesigned TOEFL is required for acceptance into most English-speaking university programs.

The new TOEFL has four sections, each of which tests one of the four major language skills: listening, reading, writing, and speaking. A much-touted feature of the new iBT test is that it also involves the integration of various skill types; for instance, test-takers are required to

read and hear information about a topic and then speak or write a response to it. ETS claims that these tasks are representative of the skills that will be needed in a university program (TOEFL, 2007b). To increase the authenticity of the experience, the scripts used in the listening section are taken from interactions in classrooms, labs, study groups, and administrative situations (such as the registrar's office), while the written texts are taken from textbooks and course materials.

The newly overhauled test contains a variety of question types that expand on the older version's multiple-choice format. In the reading section, learners read a passage written in academic style and answer a variety of questions about it. Some are basic content questions, referring to the main idea of the text or details contained in it. Others require an understanding of text organization, requiring learners to fill out a chart with information from the text, or to indicate the appropriate place for a new sentence to be inserted. In the listening section, learners hear two conversations and four lectures; then they answer questions about content, implications made by the speakers, or attitude of the speakers. Test-takers are encouraged to take notes during the listening passages, as they would be able to in a classroom.

The writing section and the speaking section have both independent and integrated portions. For the independent questions, learners respond to prompts based on their own opinions and personal experience. For instance, a possible question would ask if students think that it is important to always tell the truth (TOEFL, 2007). On the other hand, the integrated questions require students to perform tasks that require multiple language skills. For instance, students read a short passage, listen to a recorded lecture, and then respond with an essay comparing and contrasting the two or summarizing the main points of both. The purpose of these integrated tasks is to better mimic the kinds of skills that are needed in academic classes.

Indeed, anecdotal evidence indicates that these are also the sections that challenge and intimidate students the most (Lois Wilson, personal communication).

The writing and speaking sections of the test are graded by a human scorer using a rubric. The rubric for the speaking test gives descriptions of the type of speech required for 6 levels of accomplishment. Speech is graded on topic development, including cohesion of ideas, support of ideas, and overall coherence; language use, including grammatical accuracy, grammatical range, and vocabulary use; and delivery, including pronunciation, intonation, and fluidity. Writing is graded from a scale of 1 through 5 on a fairly holistic basis, taking into account the development of ideas, language use, and organization.

It is interesting to note that in the TOEFL-iBT, the writing and speaking sections are the only ones to make specific reference to grammar skills, and only as part of a rubric. This choice was made for two reasons. First, the test-makers wanted the test to closely mirror the kinds of tasks that students need to perform in an academic setting; obviously, discrete-point grammar tasks are not a part of most content academic university courses. Second, test-makers hope that students will benefit from positive “washback”; that is, in order to prepare for the test, they will not spend their time drilling grammar, but rather developing their communicative language skills (TOEFL, 2007).

The focus of ETS’s other major ESL test, the TOEIC, is business English. Like the TOEFL, the test consists of four sections: listening, reading, speaking, and writing, with both independent tasks and integrated tasks. The tasks are geared towards business or administrative topics; for instance, a learner might have to read an email with questions about scheduling and send a reply (TOEIC, 2007). The TOEIC is slightly different from the TOEFL in that the

listening and reading sections have not been substantially updated since 1979, and they contain items that focus more explicitly on grammar (Stoynoff & Chapelle, 2005).

Interestingly, the TOEIC specifically advertises itself as a placement test for English language programs. There is no reason to believe that the TOEIC should be superior to the TOEFL as a placement measure, especially since many features of the TOEIC have not been updated for many years; neither does ETS provide any evidence to this effect. It may well be a matter of practicality; that is, the TOEIC costs less and has wider availability.

Another large-scale high-stakes test, The International English Language Testing System (IELTS) from the University of Cambridge, was originally targeted at students who planned to work or study in Great Britain or Australia, but it is increasingly being used in the U.S. (Stoynoff & Chapelle, 2005). There are two tracks to IELTS; students may choose either the academic track, which focuses on academic reading and writing tasks, or the “general training” track, which uses more general-interest topics, such as advertisements or descriptions of cities. In general, the level of proficiency expected and the types of skills that are tested are the same for both tracks (IELTS, 2007). As with the TOEFL, the listening and reading sections are machine scored, while the speaking and writing sections are hand scored. In the listening section, test-takers hear six recordings; then they answer multiple-choice questions about the main idea or details of the monolog or conversation. They may also be required to fill in information on charts or maps. In the reading section, students read three passages and answer discrete-point questions about information in the texts. For the writing and speaking sections, students produce a response to a prompt which is scored by a human grader using a rubric. The rubric for the speaking section refers to pronunciation, overall grammatical complexity and accuracy, fluency, and breadth of vocabulary, while the rubric for the writing section refers to

coherence and cohesion, vocabulary use, overall grammatical range and accuracy, and content (how well the question was answered). The rubrics are similar to those used to score the ETS exams.

The Cambridge and ETS tests represent a current understanding of the nature of communicative language proficiency. They attempt to directly measure the specific language skills that will be required in a work or academic setting, and as such, are well designed. Although they are not generally marketed as such, some schools might consider using them as placement tests, since they are intended to measure communicative competence. However, the drawbacks of these tests are lack of convenience for students and high cost. The TOEFL-iBT costs \$140 as of April 2007, while the IELTS costs at least 170 Euros. Tests must be scheduled in advance and can only be taken at approved locations, of which there are relatively few, meaning that test-takers may need to drive or fly to a testing site. The TOEIC is cheaper at \$65 and may be more widely available², but it is still potentially problematic for many English programs to use on a regular basis. For this reason, these tests are generally not suitable to use as a placement test for most intensive English programs.

2.1.2.2 Other proficiency and placement measures

The University of Michigan produces a number of English language proficiency tests, including The Michigan Test of English Language Proficiency (MTELP) and Michigan English Language Assessment Battery (MELAB). These tests are more traditional than those produced by the University of Cambridge and the Educational Testing Service, in that they contain sections

² An institutional form of the TOEFL and the TOEIC can be purchased, which reduces the price.

testing grammar and vocabulary directly with multiple choice questions in a “Grammar/Cloze/Vocabulary/Reading” section. Examples of a grammar question and a vocabulary question are given in Table 1.

Table 1. MTELP grammar and vocabulary questions (Michigan, 2007)

1.	- How did you find out about the surprise party?
	- Janet told me, although she _____ .
	A. mightn't have
	C. won't have
	B. ought not to have
	D. mustn't have
2.	Sally's room was such a mess. I have never seen anything so totally _____.
	A. impeccable
	C. cluttered
	B. immaculate
	D. sumptuous

Stoyhoff & Chapelle (2005) point out that these types of questions do not have the qualities of authenticity and interactivensess that are thought to be the hallmark of communicative language testing. For instance, there is no effort to keep topics consistent throughout the test taking experience; instead, the questions move from topic to topic in random fashion. Another drawback to the test is that grammar is scored together with vocabulary and reading ability. The test designers do not clarify why they do not separate scores on these measures, but the test would have greater potential as a placement test if these scores were kept separate, especially for programs that have separate classes for grammar and reading.

The listening portion of the Michigan tests is delivered by a recording which presents short utterances and longer 3-4 minute discourse chunks. For the short-utterance tasks, learners hear a small piece of conversation and then are asked to determine an appropriate interpretation

of the meaning or implication of an utterance by selecting the correct answer to a multiple-choice question. As in the other sections, the topics change from question to question. For the longer discourse tasks, learners hear a mini-lecture or short conversation and then answer questions about the main idea and details of the passage, as well as implications made by the speakers and the vocabulary used.

There is also a writing portion to some Michigan Tests, including the MELAB. For this part of the exam, students compose a 200- to 300-word essay to defend a position on a topic using evidence from personal experience. Essays are scored by two raters using a holistic rubric that describes the level of organization, topic development, and use of grammar and vocabulary for 10 levels.

Although the Michigan Tests may have shortcomings in terms of their authenticity and communicativeness, they do have the advantage of being relatively inexpensive, quick to take (an hour and a half or less) and easy to administer. For this reason, they are used as placement tests in a number of intensive English programs.

Similar to the Michigan Tests are the ESL Computer Adaptive Placement Exam (ESL-CAPE) of Brigham Young University, the Comprehensive Adult Student Assessment System (CASAS) (CASAS, 2007), the Oxford University Quick Placement Test (Oxford, 2007) and the Act ESL Computer Adaptive Placement test (ACT-ESL, 2007). These four tests consist of a grammar/usage section, a reading section, and a listening section. Typically, grammar/usage questions consist of multiple choice cloze-type activities or error corrections, while the reading sections present short passages and require learners to answer multiple-choice questions about the main idea and details contained in the passage. The listening sections require students to identify pictures that correspond to statements, to choose the correct summary of a piece of

spoken text, or to make inferences about speakers' intentions or attitudes. These types of tests are convenient to administer and relatively inexpensive, but they lack speaking or writing components.

Tests like these may also be criticized by those who advocate the use of authentic, interactive, communicative language testing. Clahsen (1985) suggests that the discrete-point, written exams that are generally used to measure second language ability do not give an accurate representation of learners' abilities. The first reason is that standardized language tests fail to provide an authentic environment for the expression of language, and therefore do not represent learners' authentic abilities. According to Clahsen, another reason that multiple choice tests are inadequate is that some learners, particularly those who are less educated, are completely unfamiliar with the tasks required of them in these tests and are at a large disadvantage when taking them.

2.1.2.3 The use of rubrics

All of the proficiency tests discussed so far use rubrics to score speaking and writing sections (if these sections are available). Rubrics are descriptions of language behavior that is expected at different levels of proficiency. They may be oriented to the learner, to the assessor, or to specific constructs that are expected at each level. The most common type of rubric, a learner-centered rubric, describes tasks that a learner is able to perform. An example can be found in ACTFL (American Council of Teachers of Foreign Languages) proficiency guidelines, a ten-level scale that describes proficiency for speaking, writing, listening, and reading. The descriptors for Intermediate (Level 2) speaking includes this statement: "Can ask and answer questions, initiate and respond to simple statements, and maintain face-to-face conversation" (ACTFL, 1999). Learner-centered scales may also describe limitations of the learner's language skills. For

example, the same ACTFL Intermediate speaking descriptor continues: “[the learner maintains conversation] in a highly restricted manner and with much linguistic inaccuracy.”

An assessor-centered rubric focuses on the person who interacts with the non-native speaker; in other words, it focuses on the ease with which the non-native speaker can be understood. ACTFL guidelines include some assessor-oriented statements, for instance: “Although misunderstandings still arise, the Intermediate-Mid speaker can generally be understood by sympathetic interlocutors.” Limitations in understanding are also noted: “Repetition may still be required [in order for the interlocutor to understand the speech]” (ACTFL, 1999).

Finally, construct-oriented rubrics refer to specific forms that the learner produces (or fails to produce). A hypothetical construct-oriented measure is given in (1):

- (1) The learner can use modal forms such as *can*, *could*, *will*, or *would* to express requests in a polite way.

None of the rubrics discussed here so far refer to specific grammatical forms; rather, general descriptions of learners’ ability or areas of L2 knowledge are described. One of the main goals of this study is to explore the possibility of building construct-oriented descriptors into widely-used proficiency/placement rubrics.

A number of researchers have expressed significant misgivings with rating scales such as these. Much of the debate has centered on the ACTFL speaking guidelines, a proficiency scale developed by the American Council of Teachers of Foreign Languages in 1982 and revised in 1999. The scale is used to rate a speech sample, which is elicited through an Oral Proficiency Interview (OPI) by a trained interviewer who poses questions of varying levels of difficulty. The guidelines provide descriptions for the linguistic behavior of a speaker at four levels (Novice,

Intermediate, Advanced, Superior), where each level is also broken down into Low, Intermediate, and High ratings—except Superior, which is considered the highest possible accomplishment by a non-native speaker. To give an example of an ACTFL description, a sample from the Advanced-High level is given below:

Speakers at the Advanced-High level perform all Advanced-level tasks with linguistic ease, confidence and competence. They are able to consistently explain in detail and narrate fully and accurately in all time frames. In addition, Advanced-High speakers handle the tasks pertaining to the Superior level but cannot sustain performance at that level across a variety of topics. They can provide a structured argument to support their opinions, and they may construct hypotheses, but patterns of error appear. They can discuss some topics abstractly, especially those relating to their particular interests and special fields of expertise, but in general, they are more comfortable discussing a variety of topics concretely. (p. 9)

Note that some of the descriptors are relative; that is, they refer to general “Advanced level” tasks and the tasks of the next level of achievement, Superior.

The major criticism that has been leveled at the ACTFL scale, and others like it, is that it has no theoretical or empirical basis (Lantolf & Frawley, 1985, 1988; Pienemann, Johnston, & Brindley, 1988; Young-Scholten, Ijuin, & Vainikka, 2005). That is, there is no evidence either from linguistic theory or from second language research that language acquisition occurs in the stepwise fashion described by the Guidelines. Even worse, it has been pointed out that the ACTFL OPI measures content, form, and sociolinguistic ability at the same time, but there is no

supporting evidence that these three areas can be measured simultaneously as part of a single holistic score (Douglas, 1988; Bachman, 1988).

A related problem is that the scale as a measure of proficiency is inherently circular. For instance, a student is considered an “Intermediate” learner because she can perform tasks that test designers assign to this level; however, the designation “Intermediate” has no meaning outside of that system. This circularity seems to reinforce Vollmer’s (1981) and Pienemann’s (1985) viewpoint that the only possible definition of proficiency is “a student’s performance on a proficiency test.” A large part of the problem is that many of the descriptors in the scale are relative; that is, they only have meaning in relation to other descriptive words in the scale. For this reason, Young-Scholten, Ijuin, & Vainikka (2005) argue that raters who use rubrics which describe language only in relative terms cannot be consistent, which means that students will frequently be placed into levels that are too difficult or too easy for them. One reason to include construct-oriented measures is to eliminate some of this vagueness and to provide “real world” reference points for descriptors.

Another criticism concerns the claim that the Oral Proficiency Interviews are authentic in that they replicate the type of communication that occurs in the real world—that is, there is a speaker, a listener and a message to be conveyed. However, many researchers have pointed out that the OPI is hardly the same as a natural conversation; both participants are there for a very specific purpose. Even if the linguistic tasks are disguised as casual questions, the learner is all too aware that the “conversation” is a test (Young, 1995). As Lantolf & Frawley (1988, p. 183) put it, “There is only one task in OP testing—the test”.

Finally, some researchers have criticized the OPI for focusing too much on grammatical accuracy. For instance, Bachman & Savignon (1986) argue that the Guidelines do not reflect a

current understanding of communicative proficiency, including sociolinguistic competence. As such, the OPI rating is representative of only a part of a learner's linguistic abilities. However, Magnan (2002) defends the OPI by showing that grammar is only a part of the criteria that raters use to judge learners' production. While there is a correlation between grammatical accuracy and proficiency score, there are many other factors that can "override" grammar to raise or lower the score.

2.1.3 An alternative to traditional tests: Developmental approaches

Upshur (1971) claims that second language research and language testing have an inherent connection. SLA researchers often use language tests to measure learners' progress, while testing specialists look towards SLA research for help in the design of better language tests. For example, SLA research about the effect of individual differences in second language development has led to investigations about how these cognitive differences can affect learners' performance on language tests.

However, there may be more to be gained from a stronger connection between the two fields. Some test developers feel that they have not had much support from second language research. As Oller (1991, p. 9) writes, "Language testing has not been well served by applied linguistic theory, and has been forced to reach its own solutions and compromises, or to make its own mistakes". Oller goes on to point out that some of the greatest advances in defining the notion of proficiency have been made by language testing researchers, not theorists.

Some SLA researchers do want to have a stronger impact in language testing, and complain that language testers have not taken notice of findings in applied linguistics. For instance, Pienemann, Johnston, & Brindley (1988) argue that proficiency tests need to have more

of a basis in theoretical second language research. On the same point, Clahsen (1985) claims that certain theoretical findings have not been integrated in test design. Specifically, he criticizes language tests that simply calculate a learner's deviations from the target language norm by measuring accuracy of forms or number of errors, citing findings which indicate that certain grammatical errors may persist even in learners with very high proficiency (Pienemann & Johnston, 1987). Similarly, Young-Scholten, Ijuin, & Vainikka (2006) argue that current proficiency tests fail to take note of second language research which demonstrates that learners make errors while "experimenting" with language. They argue that a measure should be developed which can capture learners' creative use of language, even if it differs from target forms.

Clahsen (1985) suggests that a new proficiency test could be developed that takes second language research into account. First, such a test should rely primarily on spontaneous oral speech production. There are several reasons for making this claim. First, oral language skills are very important for most learners' communicative needs; and second, in Clahsen's view oral language is more representative of learners' true capabilities, since speech is less subject to revision, correction, etc. In addition, oral production data can be elicited in a natural way, i.e., through conversation and interaction. Tests that force oral production in an artificial way, such as with the use of picture descriptions, are likely to be seen as ridiculous by at least some adult learners, especially those who are unfamiliar with formal testing. Finally, by not overly constraining the form of learners' responses, tests that rely on spontaneous oral production allow an observation of the full extent of learners' capabilities. If desired, a test-giver will be able to create a complete description of a learner's interlanguage.

Clahsen also believes that morphosyntax should be central to the assessment of second language proficiency. To some extent, he claims, this fact simply reflects the current state of research: that is, a good deal is known about L2 morphosyntax, and a clear path of morphosyntactic development can be outlined. (Here, Clahsen is referring to his own work with Jurgen Meisel and Manfred Pienemann, discussed in section 2.2.) As no one has yet been able to describe a clear developmental path in other areas of language, such as semantics or pragmatics, morphosyntax is the only clear indicator of a learner's level of proficiency at the present time. More importantly, morphosyntax is the "structural frame" of a language, and thus central to language use.

Perhaps the most innovative aspect of Clahsen's (1985) article is the proposal to use learners' morphosyntactic development to assess their level of language proficiency. To determine a learner's level of development, Clahsen advocates using profiling, that is, building a detailed description of the learner's linguistic system by observing spontaneous oral production. The profile can be used to determine particular instructional goals for a student.

Clahsen (1985) argues that it is crucial to base a proficiency measure on linguistic development, as opposed to accuracy, for several reasons. First, proficiency measures should be based on sound linguistic research as opposed to "arbitrary" accuracy measures. Importantly, developmental measures avoid the trap of simply measuring the number of errors in learners' speech—errors which may vary in number depending on task. Secondly, developmental stages outlined by processing research are "infallible" according to Clahsen, in the sense that learners cannot skip stages or revert to previous stages. Finally, Clahsen argues, developmental measures function independently of first language, learner type, or learning environment; therefore, tests based on development do not bias against any particular group of learners.

To date, two proficiency/placement measures based on learner development have been designed: Vainikka & Young-Scholten's Organic Grammar and Pienemann's Rapid Profile. The following sections outline the theory behind each test and describe the procedures used in each one. Section 2.2 begins with the theory behind Organic Grammar.

2.2 ORGANIC GRAMMAR

2.2.1 The basis of Organic Grammar: Generative syntax

The theory upon which Vainikka & Young-Scholten's Organic Grammar is based is Generative syntactic theory. A formal theory of morphosyntax is necessary to create a morphosyntactic proficiency measure, because descriptive grammars are insufficiently fine-grained and do not permit predictors for development. Therefore, any proficiency test that aims to include morphosyntactic development must take formal theories seriously. This section presents those aspects of Generative syntax which are included in Organic Grammar, specifically lexical and functional categories (especially Inflection and Complementizers).

2.2.1.1 Lexical categories and functional categories

It has been claimed that all full-fledged languages have both lexical and functional categories (Croft, 1990). Lexical categories, such as nouns, verbs, adjectives and adverbs, are thought to have semantic content, providing for the principle meaning of a sentence (Bolinger, 1975). Lexical elements also have the important ability to assign theta roles to arguments. For instance, in the sentence *Mary pushed her*, the verb *push* assigns a semantic role of "agent" to *Mary* and

the semantic role of “theme” to *her*. Generally, lexical categories are open classes, meaning that they admit new members freely. They are often considered to be linguistically universal, meaning that all languages have them, in roughly the same way.³

On the other hand, functional categories are thought to consist of grammatical, non-theta-role-assigning elements (Grimshaw, 1990). Functional categories are closed classes; in other words, they do not admit new members freely. Examples are complementizers, determiners, tense, agreement, negation, aspect, and number. Functional categories are not thought to be universal, but rather subject to cross-linguistic variation. In fact, some researchers consider functional categories to be the sole origin of the idiosyncratic morphological behavior evidenced in individual languages (c.f. the Fundamental Parameterization Hypothesis, Chomsky 1989).

There is a general consensus that lexical categories are simpler to acquire than functional categories. This seems to be the case in first language acquisition; children appear to acquire and process lexical categories more quickly than functional categories (e.g., Brown & Fraser, 1963; Radford, 1990). Similarly, Pinker (1989) and Stromswold (1994) show that children invent neologisms for lexical categories (i.e., new nouns, adjectives, or verbs), but they never seem to invent new functional items such as tense or agreement markers.

Second language learners also seem to have more success learning lexical categories and may process them more easily (e.g., Johnson & Newport, 1986; Morgan, Meier & Newport, 1989). Given the theory that lexical categories are universal while functional categories are subject to cross-linguistic variation, it might be logical to conclude that a learner would have “access” only to lexical categories when learning an L2. Indeed, this is Vainikka & Young-Scholten’s claim in the Minimal Trees Hypothesis (1994, 1996a,b; 1998).

³ This assertion is controversial. Note that lexicalization of concepts varies cross-linguistically.

2.2.1.2 Functional categories: IP and CP

The most well-researched functional categories are Inflectional Phrase (IP) and Complementizer Phrase (CP). Because targetlike negation, modals, and questions require IP and CP, researchers often focus on these structures in addition to basic word order. This section gives a brief description of each and outlines the English constructions in which each functional category is involved.

The Inflectional Phrase is a functional projection over the lexical projection VP. In English, it is assumed that modals are base-generated as the head of IP, while auxiliaries raise to I from the verb phrase. Examples are given below; note that structures or movement not relevant to the current discussion are not fully labeled.

First, consider a clause with a modal. Modals are base-generated in IP above VP. The sentence in (2) is diagrammed in Figure 3.

(2) [_{IP} you [_{I'} must [_{VP} try this bisque]]].

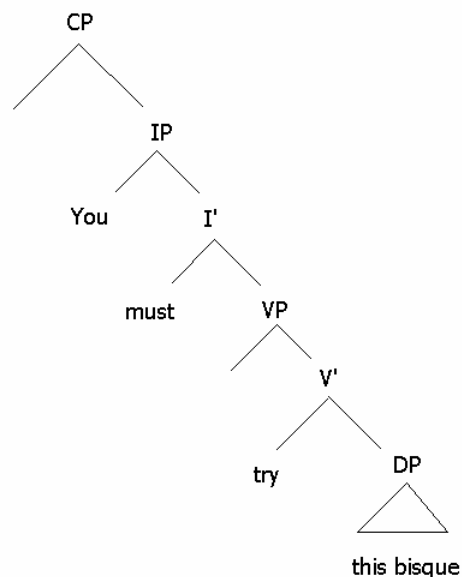


Figure 3. A sentence with a modal

Auxiliaries are also in the inflectional phrase, as shown in (3) and (4).

(3) [_{IP} you [_{I'} are_i [_{VP} e_i trying the bisque]]].

(4) [_{IP} you [_{I'} do not [_{VP} try hard]]].

Subjects also involve IP. Since Kuroda (1988), it has been assumed that subjects begin in the specifier position of VP, where they receive their theta roles, and then raise to the specifier position of IP to check inflection and receive case. The structure appears as shown in (5) and figure 4:

(5) [_{IP} you_i [_{VP} e_i took my bisque.]]

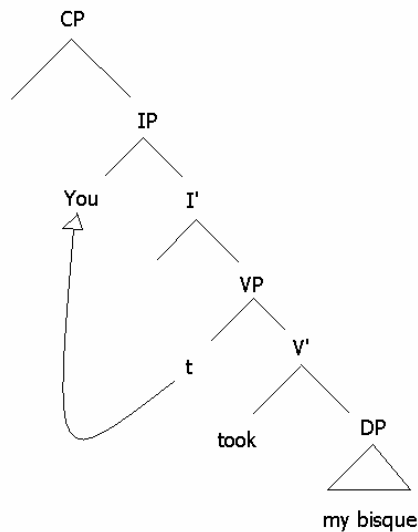


Figure 4. Subject in SPEC of IP

It has been claimed that both first and second language learners may have structures in which the subject remains in-situ, i.e., within the VP (e.g., Deprez & Pierce, 1993; Vainikka & Young-Scholten, 1998).

Tense, agreement, and aspectual marking are also all thought to be located in IP. Pollock (1989) suggested that these elements be “split” into their respective nodes: e.g., Tense Phrase

(TP), Agreement Phrase (AgrP), and Aspect Phrase (AspP). These elements may be realized in overt morphology in different ways cross-linguistically, e.g., with agglutinative affixes or fusion.

Complementizers make it possible for a clause to serve as a complement to another element in a sentence. In noun (“complement”) clauses, the head of CP is generally filled with a complementizer such as *that*, *if*, *whether*, *what*, etc. Note that different kinds of complementizers select different kinds of complement clauses, either non-finite or finite. Some examples are given in (6), (7), and (8) below.

(6) She knew [_{CP} [_C that] he liked scallops].

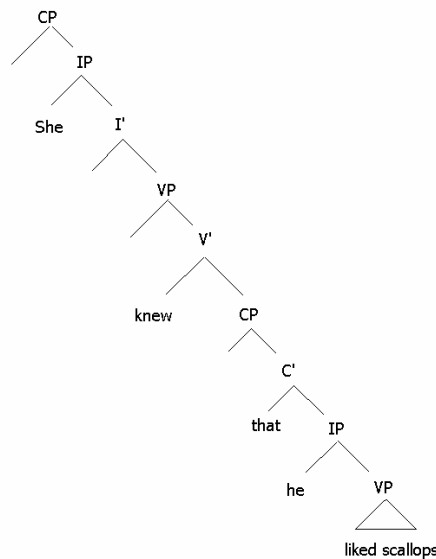


Figure 5. Sentence with CP complement

(7) He wondered [_{CP} [_C if] she would buy some lobster].

(8) I don't know [_{CP} [_C whether] to order salmon or tuna].

Adverbial clauses have essentially the same structure, as shown in (9).

(9) I ate the sandwich [_{CP} [_C because] I was hungry.]

CP is also required for so-called I-to-C movement, or subject-auxiliary inversion, which is used in English questions. Traditionally, these sentences are formed by moving the auxiliary verb by from its position in the head of IP to the head of CP. An example of yes/no questions and a comparable declarative sentence are given in (10a) and (10b).

(10a) $[_{CP}[_{IP} \text{You}[_{I'} \text{will}][_{VP} \text{buy a new hat}]]]$.

(10b) $[_{CP}[_{C} \text{will}_j][_{IP} \text{you} [_{I'} e_j][_{VP} \text{buy a new hat}]]]]?$

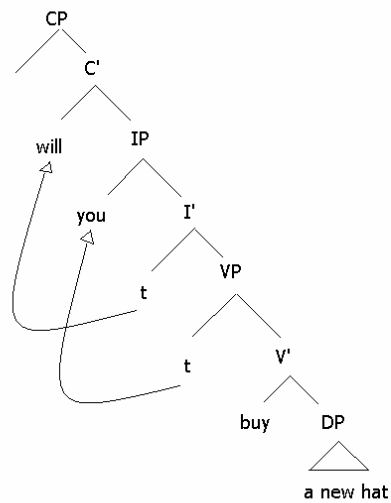


Figure 6. A yes/no question

Note that the same I-to-C movement occurs in *Wh*-questions. Under standard assumptions, there is an additional movement here: the *wh*-word moves from its DS-position to the Spec of CP. (11) gives an example.

(11) $[_{CP} \text{What}_i [_C \text{will}_j] [_{IP} \text{you} [_{I'} e_j] [_{VP} \text{buy } e_i]]]]?$

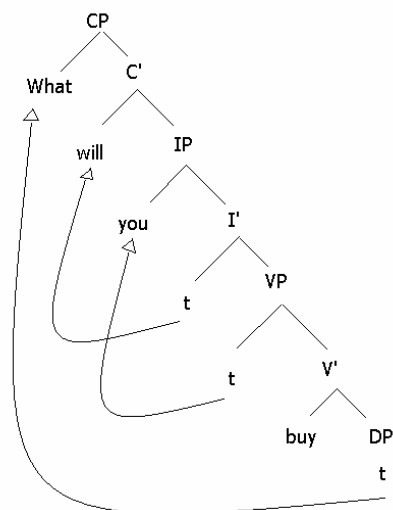


Figure 7. A wh-question

CP is also required for relative clauses. Traditionally, relative clauses are similar to *wh*-questions in that the operator (in this case, a relative pronoun) moves into the Specifier of CP position, as shown in (12).⁴

(12) I ate $[_{DP} \text{the sandwich } [_{CP} \text{that}_j [_{IP} \text{I bought } e_j \text{ yesterday}]]]$.

⁴ Rizzi (1997) has suggested that Complementizer phrase can be “split” into Topic Phrase and Focus Phrase, since there are fairly strict ordering rules for these elements. Topic and Focus phrases are involved in constructions such as the one in (13):

(13) $[_{TOPP} \text{bagels}_i [_{IP} \text{I like } e_i]]$.

As they are relatively rare in the second language data in this study, topic and focus constructions will not be central to this study.

2.2.2 The origin of Organic Grammar: L1 acquisition research

Vainikka & Young-Scholten's Organic Grammar has its origins in first language research, which was in the middle of a heated debate in the early 1990s. The main question concerned children's first language morphosyntactic development. Three theories emerged: the Strong Continuity Hypothesis, the Weak Continuity Hypothesis, and the Maturation Hypothesis. The Strong Continuity Hypothesis (e.g., Weissenborn, 1990; Boser, Lust, Santelmann & Whitman, 1992; Roeper, 1992; Schaeffer & Matthewson, 2005) proposes that children possess all functional projections at the start of language learning—in other words, there is no fundamental difference between the child grammar and the adult grammar in this domain. According to this hypothesis, the reason that children do not speak like adults is that it takes time for them to learn lexical items and to identify the particular features and the strength of those features in the language.

On the other hand, the Maturation Hypothesis (e.g., Borer & Wexler, 1987; Radford, 1990; Newport, 1990; Felix, 1992) claims that children do not have all of the functional categories in their grammar that adults do. Instead, various properties of UG emerge over time when biological constraints allow them.

A third approach is the Weak Continuity Hypothesis (e.g., Pinker, 1984; Clahsen, Eisenbeiss & Vainikka, 1994; Clahsen, Eisenbeiss, & Penke, 1996; Marinis, 2003), which is similar to the Maturation hypothesis in its claim that children lack some of the functional categories that are available in adult grammars; however, maturational factors are not considered central to language development. According to this proposal, children have access to Universal Grammar, which shapes their development. Functional categories emerge not due to biological changes, but rather because lexical learning drives the postulation of new, UG-constrained categories. For instance, the acquisition of agreement markers may lead children to develop an

Agreement Phrase (or Inflectional Phrase). Crucially, supporters of the Weak Continuity hypothesis assume that morphological marking is direct evidence of the presence of a corresponding functional category in the grammar: for instance, use of inflection indicates an Inflectional Phrase, while use of number marking indicates Number Phrase (or a similar functional category).

Vainikka & Young-Scholten, supporters of the Weak Continuity hypothesis for first language acquisition, believe that it can also be applied to second language acquisition (Vainikka & Young-Scholten, 1994). That is, they argue that L2 learners lack certain functional categories at the start of language learning; that lexical learning (i.e., identification of functional heads) drives the postulation of functional categories in L2 acquisition; and that morphological marking can be taken as direct evidence of the presence of associated functional categories.

However, in order to account for L2 data, Vainikka & Young-Scholten need to modify the framework of Weak Continuity somewhat. The result is the Minimal Trees (MT) hypothesis. The main claims of MT are as follows:

- The initial state of L2 learners is the L1 “bare VP”, that is, a lexical projection which is transferred from the first language grammar.
- Learners progress through at least three distinct stages when acquiring an L2. Each of these stages is characterized by the presence (and absence) of certain functional categories.
- Learners acquire functional categories when they identify functional heads in input. There is no first language influence in the acquisition of functional heads; therefore, after the initial state, L2 acquisition should proceed similarly to L1 acquisition.

Vainikka & Young-Scholten provide evidence for these claims and articulate these ideas more fully in a series of articles (Vainikka & Young Scholten 1994; 1996a, 1996b, 1998, Young-Scholten, Ijuin, & Vainikka, 2005; Young-Scholten & Ijuin, 2006). The following sections review these articles and provide both supporting evidence and counterevidence to the Minimal Trees hypothesis.

2.2.3 Minimal Trees Theory

Most of the evidence for the Minimal Trees hypothesis comes from work done by Vainikka and Young-Scholten with spontaneous and elicited production data from a group of naturalistic learners of German. The following sections will review the main claims of the MT hypothesis, as well as supporting evidence and counterclaims.

2.2.3.1 The L2 Initial State: A transferred VP

MT proposes that learners at the first stage of second language acquisition will transfer the properties of lexical categories from their native language. Importantly, this prediction includes the transfer of VP-headedness from the native language. In order to support these claims, Vainikka & Young-Scholten (1994) examine oral production data from 11 Turkish- and 6 Korean-speaking learners of German at mixed levels, to determine the headedness of their clauses at the initial state. Turkish, Korean, and German all have head-final VPs, as illustrated in (14) through (16).

- (14) Korean (Vainikka & Young-Scholten, 1994, p. 270)

Peter-nun umsik-ul mok – kosip – cian – ass – ta

Peter+TOP food + ACC would-like +NEG+PAST+DECL

‘Peter would not have wanted to eat food.’

- (15) Turkish (Vainikka & Young-Scholten, 1994, p. 268)

Bu kitab-ı ev-de oku-ma

this book + ACC home+LOC read+NEG

‘Don’t read this book at home!’

- (16) German

Sie meint, dass er ein-en neu-en Tisch hat.

She thinks that he a-ACC new-ACC table has.

‘She thinks that he has a new table.’

There is a crucial difference, however, between German and the other two languages. That is, German is a V2-language, which means that verbs in matrix clauses raise to head-initial COMP and give the appearance of a head-initial VP. See example (17) below, which is diagrammed in Figure 8.

- (17) Maria hat Freunde
 Maria has friends.

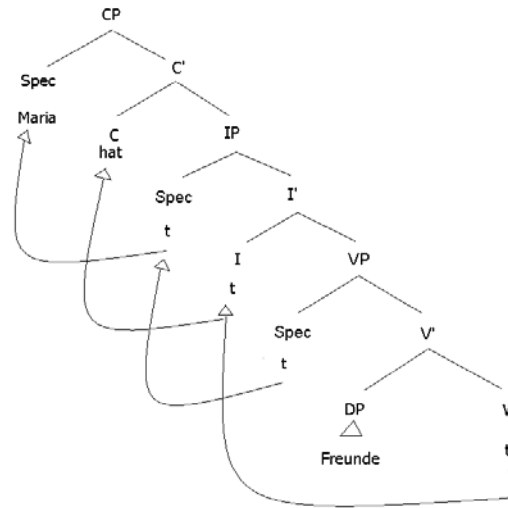


Figure 8. Movement in German matrix clauses

Because of this feature of German, it is possible that learners will mistakenly assume that the German VP is head-initial.

Minimal Trees predicts that Korean and Turkish learners of German will not make this error (at least at first); rather, they should have as their initial state a head-final VP. Indeed, V&Y-S show that the learners' utterances are consistently head-final; that is, 80% of all utterances are head-final across stages of development. More importantly, those learners who are still at the first stage of development (n = 3) had nearly 100% head-final VPs. Examples from learners at this early stage are given in (18) and (19).

- (18) Oya Zigarette trinken (Aysel #11)
 Oya cigarette drink
 'Oya smokes cigarettes.'

(19) Eine Katze Fisch alle essen

(Changsu #033)

a cat fish all eat

‘A cat ate the entire fish.’

Of course, there are other possible explanations for this behavior aside from first language transfer. For instance, it could be claimed that these learners have SOV word order because they hear clauses with final verbs and correctly recognize that German is an SOV language. Therefore, to strengthen their argument, Vainikka & Young-Scholten (1996b) examine data from Spanish- and Italian-speaking learners. This additional data is important because Spanish and Italian both have SVO word order; thus, if learners begin learning German with a head-initial VP, it can convincingly be attributed to transfer from the L1s. V&Y-S examine longitudinal data from four Italian speakers and one Spanish speaker from the ZISA corpus (Meisel, Clahsen, & Pienemann, 1981), as well as cross-sectional data from six Spanish speakers in the LEXLERN project (Clahsen, Vainikka, & Young-Scholten, 1991). They conclude that four of these 11 speakers (two Italian and two Spanish) are still in the initial stage of learning because they have mostly head-initial structures; thus it appears that they have transferred the headedness of the VPs in their first language.

This finding is consistent with previous studies. For instance, duPlessis et al. (1987) shows that English and French speakers begin with SVO order when learning German, while many studies (e.g., Clahsen & Muysken, 1986; Eubank, 1988; Schwartz & Tomaselli, 1990; Tomaselli & Schwartz, 1990) have found that Spanish, Italian and Portuguese speakers begin learning German with SVO word order. Indeed, the claim that learners transfer the headedness of their VP may be the least disputed aspect of the Minimal Trees hypothesis.

2.2.3.2 Lack of functional categories at the initial state

Another important claim of the Minimal Trees hypothesis is that only lexical categories are present in the first interlanguage stage. That is, beginning learners are predicted to lack any elements requiring a functional projection, such as auxiliary verbs, agreement markers, complementizers, and verb raising or Wh-movement⁵ (Vainikka & Young-Scholten 1998, p.21). In principle, a learner at this early stage should be able to produce noun phrases, verb phrases (with the headedness of their L1), and adjective phrases; however, V&Y-S focus their attention on the production of VPs. Vainikka & Young-Scholten (1994; 1996b, 1998) consider data from seven learners of German with varied language backgrounds who are thought to be at the VP-stage and conclude that the learners produce very few auxiliaries, consistent with the predictions of MT.

There is a wider range of behavior in regards to verbal inflection. German has a fairly rich inflectional system; as a reference, the German present-tense verbal agreement paradigm is given in Table 2.

⁵ Note that this description appears consistent with Klein & Perdue's (1997) Basic Variety. Indeed, Vainikka & Young-Scholten (2006) argue that the Basic Variety is simply a stage in the path of acquisition that they outline.

Table 2. The German subject-verb agreement paradigm for the verb *schwimmen*, ‘to swim’

	Sing	Plural
1 st person	schwimm-e (colloq: schwimm- \emptyset)	schwimm-en
2 nd person	schwimm-st (formal: schwimm-en)	schwimm-t (formal: schwimm-en)
3 rd person	schwimm-t	schwimm-en

Because agreement marking involves IP, MT predicts that learners at the early stages will lack productive agreement. Indeed, Vainikka & Young-Scholten report that the Romance speakers range from 11% accuracy to 36% accuracy in their agreement marking on verbs. Unfortunately, it is less clear how to evaluate the production data from the Korean and Turkish learners, as V&Y-S present the data in a different format. With these learners, V&Y-S assert that either the *-en* ending is used or a bare verb (no ending) appears as a default form, as is common in German children (Clahsen, 1991). Indeed, this conclusion may accurately describe the behavior of two learners: “Aysel”, who marks 92% of his verbs with *-en* and 8% with bare forms, lacking any other agreement markers, and “Memduh”, who behaves similarly. An example is given in (20).

- (20) Meine Vater nicht rauchen Sigara. (Memduh. V&Y-S 1994, p. 282)
my father not smoke cigarette
‘My father doesn’t smoke cigarettes’

However, it is not clear what conclusion to draw about a Korean learner, Changsu. V&Y-S report that Changsu uses the “default suffix” *-en* 68% of the time, the bare form of the verb 9% of the time, and other suffixes 23% of the time. Unfortunately, they neglect to report what

percentage of the “default suffixes” *-en* are actually used in appropriate contexts. Note that *-en* is the correct marker for 1st person plural, 3rd person plural, and 2nd person singular and plural formal—all contexts which are highly likely to have been present in the production data. Obviously it is not appropriate to consider all uses of *-en* as default suffixes when other endings of the paradigm are present. Despite this oversight, it is clear that there is at least a general pattern for many learners in which faulty agreement marking is used in the VP-stage, and the claim that beginning learners lack productive verbal morphology has not been widely contested.

The above arguments relate to the absence of Inflectional Phrase in learners’ grammars. MT also predicts a lack of Complementizer Phrase in early stages. Vainikka & Young-Scholten (1994; 1996a,b; 1998) argue that the seven learners in the VP stage lack embedded clauses and yes/no questions—two structures requiring CP. Although some of the learners do produce *Wh*-questions, V&Y-S argue that these *Wh*-questions can be analyzed as lacking a CP.

Some studies by other researchers have supported aspects of V&Y-S’s argument that no functional projections are available at the L2 initial state. For instance, Rule & Marsden (2006) examine elicited production data of French negatives (specifically, the form *pas*) by English-speaking learners and determine that English speakers are unable to raise verbs; they therefore conclude that no functional projection is available at very early stages of acquisition. However, contrary to the claims of MT, Rule & Marsden comment that some problems in the realization of morphology may persist even after functional categories have been acquired, consistent with the Missing Surface Inflection Hypothesis (e.g., Haznedar & Schwartz, 1997).

Another study that supports aspects of the MT hypothesis is Bhatt & Hancin-Bhatt (2002). Bhatt & Hancin-Bhatt examine elicited production data and sentence interpretation data from over 200 Hindi-speaking learners of English in schools in India. They find that learners

cannot interpret adverbials with CP and fail to produce embedded clauses or questions, supporting the claim of the Minimal Trees hypothesis that CP does not transfer to early grammars. However, contrary to MT, they conclude that early learners do have access to IP, as they use them to form small-clause-like structures.

On the other hand, Vainikka & Young-Scholten's claims regarding the initial state have met with significant criticism. Grondin & White (1996), Schwartz (1998) and Epstein, et al. (1998) argue that the lack of a particular structure in learners' production does not necessarily indicate the lack of the functional projection associated with it. To draw an extreme example, it is not expected that learners who have had a single hour of exposure to an L2 will learn enough vocabulary to produce an embedded clause, but this fact is not enough to conclude that the learner has no underlying knowledge of embedded clauses. Epstein, et al. also suggest that performance constraints may inhibit oral production of certain structures, even if underlying competence is intact; therefore, they conclude, the conclusions drawn by V&Y-S regarding the initial state are not valid. Schwartz (1998) also objects to the Minimal Trees analysis of the initial state for conceptual reasons; she argues that it is difficult to imagine what kind of cognitive principle would allow the kind of selective transfer that V&Y-S propose. This transfer would, for instance, allow a learner to process only part of the subcategorization of some verbs—i.e., the fact that the verb *want* can be followed by an NP, but not the fact that it can be followed by a functional phrase.

Most of the empirical data against MT has been taken from studies of child L2 acquisition. For instance, Lakshmanan (1993/1994, 1998) examines oral production data from Spanish, French and Chinese learners of English from ages 3-11. She notes that the copula *be* and auxiliary *be* emerge very early on in the ESL learners' production—much earlier than it does

in native speakers of English—indicating the presence of Inflectional Phrase. Similarly, negation appears appropriately, before main verbs but after modals, which is suggestive of verb-raising (and therefore, the presence of IP). Finally, children use infinitival complements, such as “I want to see you,” which requires a functional phrase. A similar study is Grondin & White’s (1996) examination of production data from two English-speaking child learners (4;5, 4;9) of French. The children produce determiners, questions, and appropriate inflection very early. Grondin & White argue that there is a gradual, quantitative difference in the children’s production of these structures, not a sudden improvement as one might expect with the acquisition of a functional category. In another study of child L2 acquisition, Haznedar (2003) argues against the MT hypothesis by showing that Erdem, a four-year-old Turkish speaker who moved to the UK, spontaneously produced the copula and questions very early.

There are also a few studies of adult L2 learners’ acquisition of functional categories. Epstein, et al. (1998) examine data from 17 low-level adult Japanese-speaking learners of English who performed a repetition task. The learners were asked to repeat a variety of structures, including modals, topicalizations, certain inflections, negation, relative clauses, and wh-questions. Although learners had difficulty, they were able to repeat many structures with IP or CP, such as modals, inflection, topicalizations, and relative clauses. Similarly, Dube (2000) argues that English-speaking learners of Zulu successfully transfer CP to their English-Zulu interlanguage, as is evidenced by learners’ early use of the complementizer *ukuthi* ‘that’ in subordinate clauses.

One counterargument to studies like those above comes from Myles (2004), who suggests that much of the evidence against the Minimal Trees hypothesis is drawn from an overestimation of the linguistic competence of the learners. Myles argues that some learners

who appear to have productive morphology or language structures may simply be employing non-productive chunks of memorized speech. Unfortunately, this issue is difficult to resolve, as it is not a simple task to decide which forms are chunked and which are productive. We will return to the issue of chunking in the Discussion section.

Another counterargument is that researchers who find evidence against MT begin collecting data after the VP stage has been passed; that is, the learners are no longer at the initial state. Gruter (2004) attempts to avoid this problem by using comprehension tasks with very early learners who are unable to produce complex language. In her study, English-speaking learners of German are able to correctly answer ambiguous *wh*-questions, possibly indicating that they can process the CP correctly.

To summarize this section, it seems likely that some of the predictions of Minimal Trees regarding the initial state are too strong, although it may be possible to adjust the theory to accommodate new data. The next section outlines further claims of the MT hypothesis.

2.2.3.3 Gradual development of phrase structure

Minimal Trees predicts that L2 learners progress through at least three stages in development. First, learners begin with a “bare VP,” that is, a Verb Phrase which they transfer from their native language. The headedness of the VP will match the VP of the native language. For instance, a speaker of a language with SOV order might produce the sentence in Figure 9.

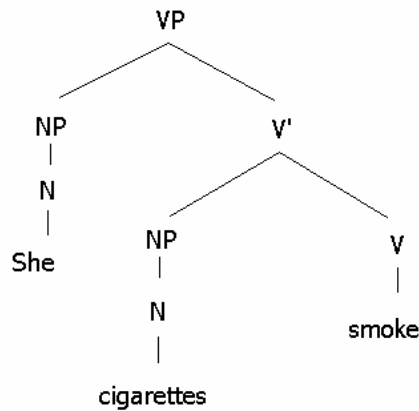


Figure 9. A learner utterance at the VP stage

According to Vainikka & Young-Scholten, learners eventually adjust the headedness of the VP to match that of the target language (if necessary). At this stage, learners will only be able to produce short phrases without auxiliaries, modals, or inflection.

At this stage, subjects are predicted to be optional. Note that in standard Generative theory, subjects originate in the VP before moving to IP for agreement checking (Koopman & Sportiche, 1991). Therefore, it is logical to predict that learners at the VP stage will often lack overt subjects, as they have no position for the subject DP (or NP for these learners) to raise to. To explain the fact that some learners produce subjects, however, V&Y-S suggest that early learners may realize subjects in-situ: that is, in the Spec-VP position as shown in Figure 9. This suggestion is consistent with many other proposals regarding in-situ subjects for child language (e.g., Deprez & Pierce, 1993).

As learning progresses, the learner develops a Functional Phrase, which provides a position for verb raising. V&Y-S argue that learners in this stage alternate between bare VP-structures and sentences with raised verbs. However, the learners have not yet acquired the

morphological agreement paradigm, because they lack Agreement Phrase. Typical utterances from this stage are given in (21) and (22), diagrammed in Figure 10 below.

(21) Ich sehen Schleier. (Kemal 604, V&Y-S 1994: p. 289)

I see veil

'I see (the) veil.'

(22) Jetzt brau Wohnungsamt fragen. (Sevinc, 111: V&Y-S 1994, p. 289)

now (I) need housing.authority ask

'Now (I) need to ask (the) housing authority.'

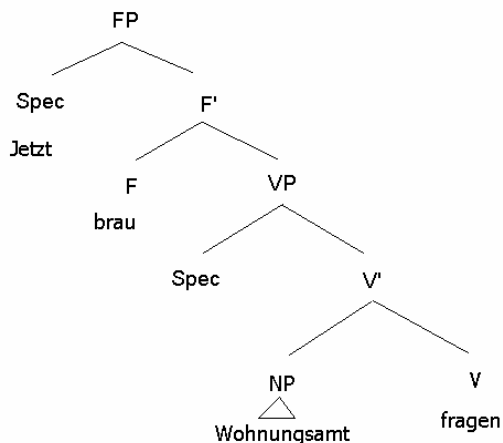


Figure 10. A learner utterance at the FP stage

According to Vainikka & Young-Scholten, learners at this stage might use the Functional Phrase to produce modals and auxiliaries; however, because IP (or Agreement Phrase (AgrP) in V&Y-S 1994) has not been acquired, inflection will be faulty. Learners at this stage will also fail to produce CP structures such as non-formulaic questions or embedded clauses.

When learners acquire the agreement paradigm for verbs, they have entered into the Agreement Phrase (AgrP) stage. As a full-fledged functional category, AgrP allows for verb-raising, modals, and inflected auxiliaries and main verbs. The development of a CP may also

begin during this stage, although production of embedded clauses and Wh-questions may be slow at first.

The proposed structure is identical to that shown for the FP stage, but AgrP replaces FP. As an example, the German utterance in (23) is diagrammed in Figure 11.

(23) Jetzt brauche ich Wohnungsamt fragen.

Now need I housing.authority ask

‘Now I need to ask the housing authority.’

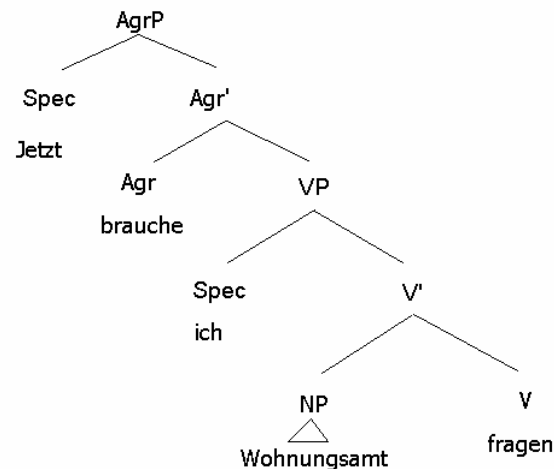


Figure 11. A learner utterance at the AgrP stage

In the final stage, learners acquire the CP because they add complementizers such as *that* or *whether* to their vocabulary. In the CP stage, learners are able to produce structures such as the wh-question in Figure (12).

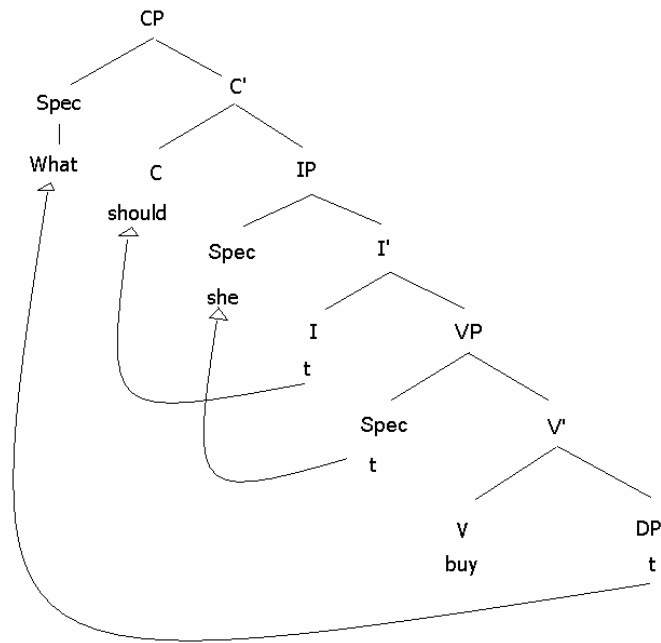


Figure 12. A learner utterance at CP (final) stage

Some researchers have disputed V&Y-S's proposal that IP will be acquired before CP, and that there will be no influence from the first language on the acquisition of these structures. For instance, Schwartz (1998) argues that MT has no way to account for a well-documented verb-raising error of French-speaking learners of English, that is, the production of sentences such as "John drinks quickly his coffee." Presumably, these errors are due to L1 influence on the acquisition of the English structure; however, MT does not allow for this influence.

Haznedar (2003) argues against the sequential development VP, IP, CP, using longitudinal development data from Erdem, a Turkish-speaking child learner of English. She points out that Erdem appears to have both IP and CP structures early on in development, and that IP structures do not precede CP structures. Furthermore, verbal inflection does not appear to correlate with the development of functional categories, but remains low even after Erdem has produced many IP and CP structures. Haznedar argues for the Missing Surface Inflection

Hypothesis, which claims a dissociation between syntax and morphological marking. There are no studies with adult L2 speakers comparable to the Haznedar (2003) study, although such a study would strengthen Haznedar's argument.

Despite the criticism of many aspects of the Minimal Trees hypothesis, Vainikka & Young-Scholten, in cooperation with Ijuin, an ESL scholar, have developed a placement measure which is based on Minimal Trees theory. This test, along with Pienemann's Rapid Profile, is one of the first placement/proficiency tests that is based on a path of morphosyntactic development. The next section details the procedures and measures used in the test.

2.2.4 A placement test based on Minimal Trees

This section presents Young-Scholten, Ijuin, & Vainikka's (2006) Organic Grammar (OG) placement test. The OG test uses written production, on the grounds that it mimics spontaneous speech while being much easier to collect from large groups than oral data. The procedure is as follows. First, learners are told that they will be evaluated on a memory task. They then view slides of a ship traveling down an overflowing river. The ship hits a bridge, flips under the water, and turns upright again on the other side of the bridge. Following these slides, a series of photos are shown of various objects and people. Learners have only a few minutes to write a brief description of the events and objects in the slides; then they view the slides again to add any information that they may have missed. The reason for the time constraints and the concealment of the true purpose of the task (i.e., linguistic assessment) is to keep the writing as spontaneous as possible, in order to reduce the influence of learners' metalinguistic grammatical knowledge about English, which the authors claim may not reflect their true acquisitional stage (c.f., Krashen, 1987).

Instructors read through learners' production (which may be as little as 50 words) to determine the presence or absence of the forms listed in Table 3. This table largely reflects the morphosyntactic development predicted by Minimal Trees. For instance, note that beginning learners are predicted to have the word order of their native language, while lacking any verbal inflection; these features are thought to be present in the VP stage of language development. Complex syntax (requiring CP) is not predicted to appear until later stages.

Table 3. Organic Grammar (Young-Scholten, Ijuin, & Vainikka 2006, p. 9)

	Word order in declaratives	Types of verbs	Verbal agreement & tense marking	Pronouns	Complex syntax
1	Initially resembles NL	Thematic (main) verbs only	None	Pronouns absent	none
2	Resembles the NL	Thematic verbs; copula 'is' appears	None	Pronoun forms begin to emerge	Formulaic or intonation-based Qs
3	Resembles the TL	Thematic verbs; modals; copula forms beyond 'is'	No agreement, some tense, some aspect, but not productive	More pronouns, but they can still be missing	Qs formulaic or w/o inversion; conjoined clauses
4	Resembles the TL	Thematic verbs, modals, copula forms beyond 'is', range of auxiliaries emerges	Productive tense, aspect, agreement with "be" forms	Pronouns obligatory, 'there' and existential 'it' emerge	Productive Qs, but may still lack inversion; simple subordination
5	Resembles the TL	Complex tense, aspect forms; passives; range of thematic verbs, modals, auxiliaries	Forms usually correct, apart from those newly attempted	Use of 'there' and 'it' beyond stock phrases	All Qs with inversion; complex subordination

Young-Scholten, Ijuin & Vainikka (2005) test the Organic Grammar system with 44 ESL students of varying L1 backgrounds enrolled in an intensive English program. They find that Organic Grammar places students in a similar manner to the more traditional method that had

been used: a reading task and writing sample. However, where the two measures differ, Organic Grammar is superior.

One of the goals of this study is to determine whether the path of development outlined in Organic Grammar can account for semi-spontaneous oral production data of ESL learners, and, if so, whether OG can be used as a placement tool. OG is not, however the only assessment tool based on L2 development. In fact, Pienemann's (2003) Rapid Profile preceded OG. The next section presents the theoretical underpinnings of Rapid Profile: Processability Theory.

2.3 PROCESSIBILITY THEORY AND RAPID PROFILE

2.3.1 The theoretical basis of Processability Theory: LFG

Pienemann (1998) uses Lexical Functional Grammar (LFG) as the basis for his processing theory. As Pienemann explains it, LFG is useful because it provides an account of three crucial procedures: the classification of grammatical information for a lexical item; the temporary storage of that grammatical information; and the operation of the information at a different position in the c-structure (constituent structure). It is ideal as the basis of a processing theory of acquisition because it is relatively simple to apply memory constraints to it, and it contains a theory of feature unification, which Pienemann claims is crucial to explaining the developmental patterns discovered in L2 German data.

There are three components to LFG. The first is constituent structure, or c-structure, which is similar in many ways to the phrase structure used in Generative grammar (described in

section 2.2). C-structure is generated directly from phrase structure rules. An example is given in Figure 13 below, which is the c-structure for (24).

(24) Jon likes these scallops.

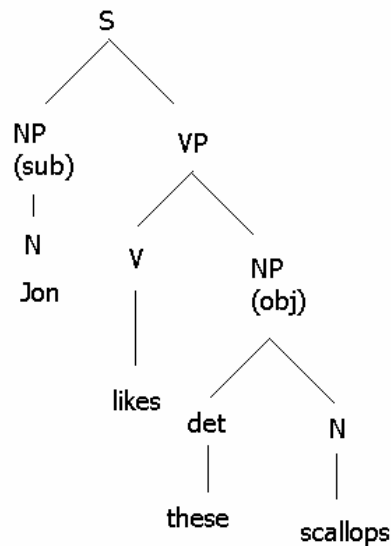


Figure 13. An example of c-structure in LFG

The second component of the linguistic system in LFG is the lexicon. Each lexical entry assigns a value to features. For example, in the above example, the noun *scallops* assigns the value PL (pl) to the NUM (number) feature, as shown in (25).

(25) scallops: N, PRED = “scallops”

NUM = PL

Some lexical entries may also require values in other functional areas. For instance, some German auxiliaries (e.g., *haben* ‘have’) require the presence of a verb which has a form beginning with *ge-* (as in the phrase *Ich habe es gesehen*, literally ‘I have it seen’).

The third and final component of the linguistic system in LFG is functional structure or f-structure, which is generated by the interaction of the c-structure and the lexicon. The f-structure contains the information which is necessary to create a semantic interpretation of the sentence and to connect the individual elements in meaningful ways. It emerges out of the interaction of the lexicon with c-structure. For instance, in f-structure the verb *likes* is related to its subject and object (i.e., *Jon* and *these scallops*).

There are various well-formedness conditions that ensure that the individual elements are related in appropriate ways to create grammatical structures. The most important well-formedness condition of LFG for Processability Theory is the Uniqueness Condition (UC), which ensures that values are compatible. For instance, let us consider what happens if a sentence such as (26) is generated:

(26) Jon likes this scallops.

In this case, the determiner *this* assigns “singular” to the NUM feature, while *scallops* assigns “plural.” Because the features clash, this utterance is eliminated as ungrammatical.

One of the key factors in Processability Theory and Rapid Profile is the distance over which the UC must operate. In some cases, there is no distance at all. For instance, the lexical entry for a past tense verb *suchte* (‘looked for’) appears as follows (from Pienemann, 1998, p. 114):

(27) *suchte*: V (PRED) = “suchte” (SUBJ) (OBJ)
(SUBJ NUM) = SG
(SUBJ PERS) = 3
(TENSE) = PAST

Note that the tense information (PAST) is included in the lexical entry; that is, it does not need to be checked or exchanged with any other element. On the other hand, the plural marking on the determiner *these* and the noun *scallops* in the above example need to be matched. This type of

matching is called a ‘phrasal procedure’ because the exchange of information occurs within a single phrase. Other procedures take place over longer distances; for instance, subject/verb agreement occurs between two phrases. How this theory relates to Processability Theory and the acquisition of English is explained in the following section.

2.3.2 The origins of Rapid Profile: Processability Theory

Manfred Pienemann’s outline of ESL development has its basis in theories of language processing that emerged in the 1970s (e.g., Bever, 1970; Forster, 1979), as well as Levelt’s 1989 model of language generation. In this work, language processing is seen as a system of computational procedures that operate on linguistic knowledge, but are largely separate from it. Because these procedures are thought to be universal, some consider them to be an ideal way to explain universal patterns in both first and second language acquisition.

The first major study to use processing as the key to explaining developmental patterns was conducted by the ZISA research group (ZISA = *Zweitspracherwerb Italienischer und Spanischer Arbeiter*, or the *Second Language Acquisition (of German) by Italian and Spanish workers*) (Clahsen, 1980; Pienemann, 1980, 1981; Meisel, Clahsen & Pienemann, 1981; Clahsen, Meisel & Pienemann, 1983). The ZISA project was a large series of studies on the second language acquisition of German, consisting of cross-sectional data from interviews with 45 Spanish and Italian speakers, as well as longitudinal data from twelve speakers. The ZISA group, who focused mostly on the development of word order, determined that acquisition of German took place in a series of stages, summarized below. This order of acquisition for German has also emerged in other work (e.g., Pienemann, 1987; Jansen, 1991).

(28) **Stage x = Canonical Order**

die kinder spielen mit ball

the children play with the ball

Stage x + 1 = Adverb preposing (i.e., beginning a sentence with an adverb.)

da kinder spielen

there children play

Stage x + 2 = Verb separation

alle kinder muss die pause machen

all children must the break have

Stage x + 3 = Inversion

dann hat sie wieder die knoch gebringt

then has she again the bone brought

Stage x + 4 = Verb Final

er sagt, dass er nach hause kommt

he said that he home comes

Importantly, the learner accumulates these rules cumulatively; that is, presence of a more advanced rule entails the existence of all less advanced rules. Additionally, the ZISA researchers claim, it is not possible for a learner to “skip” a stage. This type of pattern is known as a *strict implicational sequence*.

Clahsen (1984) argues that a simple set of processing strategies can explain the order of acquisition presented above. Expanding on principles developed by Bever (1970) and Bever & Townsend (1979), he proposes three main stages of processing development for L2 learners. Each stage consists of a strategy that the learner uses to parse L2 input and produce speech.

According to Clahsen, the strategies are necessary because second language learners' linguistic skills are not yet automatized; that is, they require a good deal of working memory space to function.⁶ As such, there is little room for complex processes, so that learners are forced to simplify linguistic operations, thereby circumventing the more thorough comprehension/production process that native speakers use. Clahsen's proposed stages of processing development are listed below in (29).

(29) 1. **Canonical Order Strategy**

Language must be processed as a fixed sequence (canonical word order).

2. **Initialization-Finalization Strategy**

In underlying sequences, [XYZ] permutations are blocked which move X between Y and Z or Z between X and Y (i.e., operations must take place at the beginning or end of a phrase).

3. **Subordinate Clause Strategy**

In subordinate clauses, permutations are avoided.

The stages outlined in (29) are interpreted differently than the descriptive developmental hierarchy described in (28), in that an earlier stage may be "cancelled" or abandoned as a learner becomes more sophisticated.

The origin of Clahsen's three strategies is as follows. Canonical Order Strategy is modified from Bever (1970), who argued for a Noun-Verb-Noun strategy of L2 speakers, based on comprehension studies. The principle underlying this strategy is simple: learners avoid complexity of structure, opting rather for a direct mapping from semantic structure into syntactic

⁶ Although a good deal of research has focused on the role of automatization and working memory in L2 acquisition, this account of L2 development has not yet been supported directly by empirical research.

strings. As Pienemann (1998) explains, these syntactic strings might be described as “flat,” in that they lack hierarchical structure.⁷

The second stage, Initialization-Finalization Strategy, is based on research into memory. That is, it has often been demonstrated that is easier to remember the first or last form in a string than internal forms (i.e., the “primacy effect” and “recency effect”). For this reason, it may be easier for learners to process forms that appear in these positions. Finally, the third stage, Subordinate Clause Strategy, was created by Clahsen (1984) to explain the finding that subordinate clauses are processed differently than main clauses. That is, learners appear to master the word order in subordinate clauses later than the word order in main clauses.

This analysis has met with some criticism. For instance, it has been pointed out (e.g., Pinker, 1984; Towell & Hawkins, 1994) that these stages *constrain* acquisition, but they do not provide a complete account for the development of interlanguage grammars. For instance, the stages are far too general to account for the many facets of grammatical acquisition. Another criticism of the theory is that it equates comprehension with production (White, 1991). That is, processing refers to operations that take place on input (comprehension), but the evidence for the theory is taken from learners’ output (production) of certain forms. Although it is commonly assumed that the two mechanisms tap into the same grammatical resources, it may be incorrect to attribute the same processing mechanisms to both production and comprehension. Note also that the ability of a certain position or structure for processing presupposes the availability of the position or structure in the first place. Thus Processability Theory attempts to be a theory of production, but cannot justly claim to be a complete theory of development.

⁷ This explanation is similar to Clahsen & Felser’s (2006) assertion that learners process a second language in a “shallow” way.

A final criticism comes from Pienemann (1998), who points out that Clahsen's strategies refer to target language "transformations." This conceptualization may not be appropriate in accounting for learners' interlanguage grammars, which may operate under different principles than the target language. In other words, the strategies commit the comparative fallacy (Bley-Vroman, 1983). Additionally, Pienemann argues, transformational grammar lacks psychological plausibility.

Pienemann's Processability Theory attempts to overcome some of these criticisms by using Lexical Functional Grammar instead of transformational grammar as the basis for the model. Additionally, he incorporates grammar more directly in the system by including a grammatical memory store which is available to the language processor and in which task-specific grammatical processing takes place (based on Levelt, 1989).

In Processability Theory, Pienemann claims that learners proceed through five stages of processing before development is complete. The stages, which I will outline briefly, are given in (30). They were created from previous work in processing combined with principles of LFG.

(30) Development of processing procedures (Pienemann, 1998; 2003; based on Levelt, 1989)

- i. lemma access
- ii. the category procedure
- iii. the phrasal procedure
- iv. the S-procedure
- v. the subordinate clause procedure (if applicable)

In the first stage, lemma access, a particular lemma (the meaning of a word plus its syntactic information) in the lexicon is activated⁸. Only single words or short strings will be produced by learners at this stage, and the learner must rely largely on non-linguistic strategies, such as gestures and facial expressions, for communication. No information about syntactic category (noun, verb, etc.) is available at this stage, so that it is impossible for the learner to parse phrases. No matching of features (due to the Uniqueness Condition in LFG) is possible at this point.

In the second stage, each lemma can be associated with categorical information (noun, verb, etc.). At this stage, learners may map words directly from conceptual structure into strings, but, because phrasal categories are not yet available, these strings are “flat”; i.e., there is no hierarchical structure. The result is that all strings will follow canonical word order. At this stage, the Uniqueness Condition can only operate on elements within a single phrasal category, since nothing else is available. For instance, featural information can be exchanged regarding plurals and possessive pronouns.

In the third stage, the category information stored with each lemma (e.g, N) can serve as the head to a phrasal category (e.g, NP). At this stage, there is now enough memory space to perform operations. However, only operations which affect structures at the beginnings and ends of sentences are possible, because these positions are salient and universal; that is, they require no language-specific processing and therefore little memory space.⁹ At this state, phrases are

⁸ Based on Levelt (1989), Pienemann uses the term ‘lemma’ to indicate the concept of each lexical item, but also the syntactic and phonological aspects associated with each.

⁹ This claim is partially based on research on the “primacy effect” and “recency effect,” the well-known ability for people to remember the first and last items in a series.

available for the exchange of featural information, so that an adverb or other element (such as an auxiliary) could be moved to the beginning of a sentence.

The fourth stage introduces sentence-internal operations; however, only operations that are “anchored” by initial or final positions will be possible, due to working memory limitations. Morphological marking may be present for the first time at this stage, but it must be local (within a phrase); for example, articles may be present, but agreement marking will not. Word order should be targetlike.

In the fifth stage, completely sentence-internal operations are possible; in fact, linguistic processing has become automatized enough to provide plenty of working memory space for all necessary operations to take place, including those that operate between phrases. At this stage, all morphological marking should be present, even when it requires relations between phrases, as does agreement marking.

In the final stage, a procedure is added that applies only to subordinate clauses. This procedure will operate differently cross-linguistically. In English, it operates on *wh*-noun clauses, as shown in (31):

(31) I wonder what they want. (cf. * I wonder what do they want)

Note that in English these clauses use *wh*-words without the usual auxiliary inversion that takes place in questions. Because this procedure requires “canceling” a procedure acquired earlier, it is thought to be particularly difficult for learners. Note that being at the final stage implies that all the other stages have been passed through; that is, there is no way to miss a step or backtrack.

This proposal is not uncontroversial. Some researchers have argued that the order of acquisition for a variety of languages fails to follow the predicted hierarchy. For instance, Alhawary (2003) demonstrates that the acquisition of noun-adjective agreement and the

acquisition of subject-verb agreement by English-speaking learners of Arabic do not follow the order of development predicted by Processability Theory. Similarly, Farley & McCollam (2004) argue that learners of Spanish produce forms in a somewhat different order than is predicted by the theory. Dewaele & Veronique (2001) also argue that Processability Theory is not adequate to account for the acquisition of gender in learners of French. Their study shows that intra-clausal gender marking is not acquired earlier than inter-clausal gender marking, contradicting the theory.

Another issue raised by researchers is that Processability Theory, while making generally correct predictions, may still be inadequate to account for many aspects of acquisition. This conclusion is reached by Glahn, Hakansson, Hammarberg, Holmen, & Hvenekilde, (2001), who analyze the production of adjectives and subordinate clauses in second language learners of Danish, Norwegian, and Swedish. Glahn, et al. conclude that their data does not directly contradict Processability; however, they argue that there are other factors which need to be considered to account for learners' development, such as discourse and conceptual factors.

In support of his proposal, Pienemann (2003) presents data from the acquisition of German, English (Johnston, 1985; Pienemann & Mackey, 1993), Italian, (DiBiase & Kawaguchi, 2002) and Japanese (Kawaguchi, 1996; Huter, 1998), which show that morphosyntactic forms appear in the predicted order.

According to Pienemann, because processing development occurs uniformly regardless of the native language of learners or the language being acquired, it is relatively simple to determine a learner's stage of acquisition by observing his or her production in the target language. The next section explores this possibility.

2.3.3 A processing-based proficiency measure

The measure that Pienemann (2003) advocates for assessing a learner's stage of development is the computer-based Rapid Profile system, developed out of Pienemann, Johnston, & Brindley's (1988) developmental measure for ESL students. During a Rapid Profile assessment, an observer takes note of those morphosyntactic features which are present in a learner's speech and those which are lacking. This information should allow the researcher to determine which developmental level the learner is currently in.

The procedure is as follows. The learner and a native English-speaking interlocutor engage in spontaneous conversation that is intended to put the learner at ease. A trained observer stationed at a computer listens to the learner's speech for the linguistic structures that are indicative of processing development (given in Table 4 below). The observer need only check an on-screen box to note an occurrence of a particular feature when it is heard. In order to determine that a learner has reached a certain stage, only the *emergence* of a structure is necessary. That is, only one example of each structure is required, as long as it is clear that it is a productive form, and not simply a memorized chunk. The reason for this decision is that the utterance of even one structure (as long as it is productive and not a "chunked" or frozen form) indicates that a learner has the processing capacity necessary to deal with that structure.¹⁰

During the profile, learner speech is not entirely spontaneous. The observer may request linguistic tasks, such as picture descriptions or information-gap activities, that elicit specific forms (e.g., questions). The purpose of these tasks is to ensure that a complete profile can be

¹⁰ It is logical to reason that learners with greater working memory should therefore proceed through the stages of development more quickly.

made of all forms. That is, the tasks ensure that learners who simply do not produce certain forms spontaneously are not categorized as early-stage learners if they are in fact capable of producing the forms. With the use of these tools, determining a learner's stage can be achieved relatively rapidly, within about ten minutes.

For obvious reasons, the original stages describing the development of German word order cannot directly apply to English, so the stages described by Processability Theory have been “translated” into concrete English structures. For instance, English Stage 3 operations (which refer to the beginnings and ends of clauses) include *do*-fronting and adverb preposing, but not the 3rd person singular *-s* marker of subject-verb agreement, which appears later.

The stages of Rapid Profile are presented below in Table 4. Stage 1 (at the bottom of the table) is the least advanced, while stage 6 (at the top of the table) is the most advanced level.

Table 4. Rapid Profile stages (Pienemann, 2003).

Stage	Phenomena	Examples
6	Cancel Aux-2 nd	I wonder what he wants
5	Neg/Aux-2 nd -? Aux-2 nd -? 3sg-s	Why didn't you tell me? Why can't she come? Why did she eat that? What will you do? Peter likes bananas.
4	Copula S (x) Wh-copula (x) V-Particle	Is she at home? Where is she? Turn it off!
3	Do-SV(O)-? Aux SV(O)-? Wh-SV(O)-? Adverb first Poss (Pronoun) Object (Pronoun)	Do he live here? Can I go home? Where she went? What you want? Today he stay here. I show you my garden. This is your pencil. Mary called him.
2	S neg V(O) SVO SVO-? -ed -ing Plural -s (noun) Poss -s (noun)	Me no live here/I don't live here. Me live here. You live here? John played. Jane going. I like cats. Pat's cat is fat.
1	Words, Formulae	Hello, Five Dock, Central How are you? What's your name? Where is X?

There are significant differences between these stages and those that appeared later in Young-Scholten, et. al's (2005) Organic Grammar,. For instance, a number of features are included in Rapid Profile but not Organic Grammar, such as object pronouns, possessive pronouns and nouns, and "Cancel Aux-2nd" (*wh*-noun clause) structures. Similarly, OG includes a number of forms and structures not included in RP: modals, relative clauses, adverb clauses, and expletive subjects, to name a few. RP also uses emergence of inflectional forms such as past tense marking, not suppliance in obligatory context, as an indicator of developmental stage; and RP

separates tense marking from agreement marking. Finally, RP relies heavily on questions as indicators of development. Because of all these differences, it is very difficult to compare the two systems directly.

Interestingly, Pienemann, Johnston, & Brindley stress that their stages of development should not be used as a proficiency measure. They write:

While it may be tempting for educational administrators or funding authorities to use the results of a quantifiable language test to stream learners into classes or to justify funding decisions, it is important to point out that the assessment procedure was not designed for these purposes. It is aimed at providing teachers with information concerning a given learner's developmental stage and hence at assisting them to make teaching learnable. (p. 240)

The concern is that the developmental test should not be used to marginalize certain learners or prevent them from receiving aid. Furthermore, it should be made clear that developmental tests are not intended to reflect learners' motivation, efforts, or aptitude. That is, the developmental profile is intended to be descriptive rather than evaluative.

However, it is not clear why the developmental measure could not quite constructively be used as a placement test, since according to Pienemann learners at each stage require a unique type of instructional intervention. Specifically, learners should be exposed to structures at the level that is one higher than their current stage (see Pienemann's Teachability/Learnability Theory, e.g., Pienemann, 1985, 1989). A logical next step is to use Rapid Profile as a placement test, since learners could be grouped according to their linguistic stage, and presented with materials that are appropriate to their level. Pienemann has already claimed that the information gleaned from this profile should be used to guide teaching and syllabus design

As I noted earlier, Rapid Profile refers to a fairly limited set of morphosyntactic elements. Some of the elements not involved in RP, including determiners, subjects, and copula forms, are contained in the Multidimensional Model, which is the subject of section 2.3.4.

2.3.4 The Multidimensional model

The developmental stages described in Processability Theory are not intended to explain all aspects of a learner's grammatical system; neither is every grammatical difference between two learners' production considered to be a maturational difference. Instead, the Multidimensional Model (Meisel, Clahsen & Pienemann, 1981; Clahsen, Meisel, & Pienemann, 1983) is intended to explain certain observed facts: first, the fact that there is a good deal of variation in learners' production even at a single acquisitional stage; and second, the fact that some language learners, even at very advanced levels, make errors in structures that they had presumably acquired at an earlier stage.

To account for these observations, Meisel, et al. (1981) propose two dimensions of grammatical features: developmental features and variational features. Developmental features are those presented in section 2.3.2 above; they are predicted to develop in stages as a learner's interlanguage develops. Variational features, on the other hand, do not develop predictably over time, but rather may be used differently by individual learners, depending on each learner's socio-psychological approach to language learning and the L2 culture.

In this model, there are two basic categories of learners: standard-oriented learners and simplifying learners. Standard-oriented learners avoid using simplifying structures when speaking the L2; rather, they aim to be grammatically accurate as early as possible. On the other hand, simplifying learners may omit certain grammatical structures from their speech if

communication is possible without them. To give an example, standard-oriented learners will have a high rate of copula suppliance (e.g., “She **is** a doctor”). Alternately, some types of standard-oriented learners who have not yet mastered copula use may avoid producing structures requiring the copula, so as to avoid making a linguistic error. On the other hand, simplifying learners do not focus as much on linguistic accuracy; they will frequently omit copulas (e.g., “She doctor”). Pienemann (1998) also predicts that some simplifying learners will fail to develop as quickly as standard-oriented learners and they may be more likely to fossilize in a certain stage of development. For instance, if a learner consistently fails to produce the copula, she will not be able to produce questions beginning with the copula (e.g., “Is that a dog?”), which is an important developmental feature. The learner will therefore be unable to progress past the current stage of development.

Learner orientation (standard-oriented or simplifying) is tied in with the degree to which a learner is “integrative.” A learner’s integrative motivations depend on her desire to learn the language and to become part of the L2 culture. To give an extreme example, a highly motivated, integrative learner may have a spouse and children who speak the L2; she may be very interested in the L2 culture and hope to learn as much as possible about it; and finally, she may require good command of the L2 in order to succeed at work. A learner who is simplifying or “segregative” may live and work with members of her own culture; she may dislike the L2 culture; and she may have been forced to come to the L2 country because of financial or family pressure. According to Meisel, et al., the first learner will make an effort to produce correct L2

speech whenever possible, while the second learner will simplify her speech to as to make it easier to process (e.g., “I go store”).¹¹

In order to determine which category learners belong to, Meisel, Clahsen and Pienemann (1981) developed a multiple-choice survey that probed for information about learners’ integrativeness. Questions revolve around each learner’s interest in the L2 culture; her plans to stay in the country; and the language used at home and at work. For example, one question asks how long the learner plans to stay in the L2 country, while another question asks whether the learner would like to take classes to improve ability in the L2.

Clahsen, et al. (1983) identify 14 variational features, including the omission of obligatory constituents such as the subject pronoun, lexical verbs, modals, auxiliaries, prepositions, and determiners. They argue that learners who simplify one of these elements will simplify all others. According to the authors, it is reasonable to measure a learner’s variational features in terms of their accuracy, while emergence is a more appropriate measure for developmental features.

Hudson (1993) criticizes the Multidimensional Model, claiming that Meisel, Clahsen & Pienemann did not have enough learners in their study and that the methodology was faulty. Hudson points out that some of the survey questions, which are intended to measure a learner’s integrativeness, do not appear legitimately tied to integrativeness at all; for instance, one question requires learners to report whether they are from a large city, medium-sized city, or

¹¹ Note the similarity to Schumann’s (1978) Acculturation Theory, which predicts that learners who are more integrated into the L2 culture will have greater success in language learning.

small town, and rates learners as more integratively motivated if they are from a large city. However, as Hudson points out, there is no a priori reason to assume this relationship. To give another example, it is not clear why learners who moved around frequently in their homeland should be more integrative than those who did not, but Meisel, et al. make this assumption in another question.

Another criticism that Hudson levels at the Multidimensional Model is that the statistics used by Meisel, Clahsen and Pienemann were faulty. First, ordinal and categorical data have been combined in the study, rendering the statistical procedures questionable. A worse problem resides in the way “non applicable” answers are included in the statistics. That is, if a learner cannot answer a question because it is “non applicable” (e.g., a question about the age of children when the learner has none), the response is scored the same as a learner who is highly segregative. That means that there is a bias to consider certain learners segregative if they have a number of “non applicable” answers. A final problem that Hudson points out is that certain of the survey questions necessarily cluster together, but these factors are treated as independent. For instance, questions relating to the amount of time spent in current employment, the amount of time living in the L2 country, and number of children are often related to the age of the learner, but these questions are treated as independent. Meisel et al.’s conclusions about social attitudes could therefore be an artifact of the learners’ ages and how that affects their production, rather than learners’ willingness to integrate into the L2 culture.

Larsen-Freeman & Long (1991) offer further criticism of the Multidimensional Model, saying that there is no theory to explain why some features should be developmental while others are variational. They express concern that the model would be unfalsifiable, especially because of the mix of emergence measures and accuracy measures.

This study will provide some additional information about the degree to which learners' motivation, contact with native speakers, and feelings towards the learning environment affect their use of variational features.

2.3.5 Summary of section 2

Organic Grammar and Rapid Profile claim to be measures of overall L2 morphosyntactic development which can be used to make decisions about appropriate instructional intervention. It is not clear, however, if either OG or RP is intended to supplant current proficiency measures. Current placement/proficiency measures do have significant flaws; for instance, multiple-choice-style tests measure accuracy in an inauthentic, noncommunicative setting, while interviews using rubrics such as the ACTFL scale are subjective and possibly inaccurate. However, traditional placement/proficiency measures have the advantage of testing for a wider variety of behaviors than OG or RP. The ACTFL scale, for instance, includes information about fluency, pronunciation, and vocabulary use, which are largely excluded from a measure such as OG or RP.

In this dissertation I propose the creation of a proficiency measure that bridges the gap between measures which are strictly grammar-oriented and those which contain little mention of grammar at all. One solution is to create an ACTFL-type rubric that includes specific morphosyntactic elements in a developmental sequence.

Unfortunately, the sequence of morphosyntactic development for L2 learners is not yet fully understood. In fact, it remains an open question as to whether there is a uniform path of morphosyntactic development for all L2 speakers. In the study described in the next section, I test both OG and RP as predictors of the emergence of morphosyntax in L2 speakers in the

English Language Institute at the University of Pittsburgh. The goal is to learn which developmental features might be included in an ACTFL-style proficiency measure.

3.0 THE STUDY

This section describes the current study, which has the goal of determining whether a placement test based on spontaneous production can include reference to specific morphosyntactic elements. Section 3.1 provides information about the English Language Institute at the University of Pittsburgh, where the study was conducted, as well as descriptions of the data, the participants, and the measures used. It also includes information about the linguistic background of the participants in the study. Section 3.2 provides the results of data analysis, including results from analysis of interviews. Section 3.3 is the Discussion, where research questions are addressed and directions of future research are discussed.

3.1 METHODOLOGY

3.1.1 The environment

This study took place at the English Language Institute (ELI) at the University of Pittsburgh. The ELI is an intensive English program which offers non-credit courses at four levels: pre-intermediate (Level 2), low intermediate (Level 3), high intermediate (Level 4), and advanced (Level 5). There are no classes for beginners. A student taking classes full time has five classes a day, four days a week, equaling 20 hours total; students may also elect to attend the school part

time if their visa status permits it. Each of the five classes focuses on a particular language skill—reading, writing, listening, speaking or grammar—but class activities often integrate multiple skills.

ELI instructors have a TESOL certificate or equivalent. They do not operate independently; rather, course content is strictly controlled by curriculum supervisors who select textbooks, write syllabi, design activities, and encourage certain teaching methods. Teachers assign homework on a regular basis and encourage students to use English outside of the classroom. They also monitor students' attendance and progress, so that problems can be dealt with early.

Students who come to the ELI usually plan to study at an American university, either for a Bachelor's degree or for postgraduate work. It is the aim of many students to receive a passing score on the TOEFL or IELTS as quickly as possible, so as to progress onto their university studies; indeed, many students in levels 4 and 5 begin standardized testing and continue until they have the desired result. Other students come to the ELI with the hope of improving their English enough to secure a desirable job in their home countries, to improve their efficacy in a current job, or to prepare themselves for life in the United States. Generally, students have had instruction in English in primary and/or secondary school but their proficiency is not yet good enough for them to be admitted into a university, secure a job, and so on.

In recent years, the student body has consisted largely of speakers of Arabic, Korean, and Chinese, although a large number of other language backgrounds are also represented, including French, Russian, Farsi, Croatian and Telugu. Generally students are between 18 and 30 years of age, but older students are not uncommon.

Student experiences in the ELI vary widely. For some students, especially those who take classes full time, the ELI becomes a kind of “home away from home”. That is, they befriend their classmates and spend much of their time participating in activities led by the ELI or by the university. Others have family circumstances which draw them away from the school; for instance, they may live off-campus with their parents, or they may have a newborn baby at home. These factors naturally influence the amount of time spent practicing English and working on assignments.

Students are placed into class levels based on their performance on three measures: the Michigan Test of English Language Proficiency (MTELP), the University of Michigan Listening Comprehension Test (LCT), and a writing sample. As described in Section 1, the MTELP and LCT are objective, multiple-choice tests. On the other hand, the writing sample is obtained by providing students with a written prompt of a general nature (e.g., *If you could travel to any place in the world, where would you go? Why would you pick this place?*). Students then have thirty minutes to write an essay in response. Samples are graded by hand by ELI instructors using a rubric that measures content, organization, and grammatical accuracy. Table 5 below shows the scores required to be placed into each of the four proficiency levels.

Table 5. New student level placement

ELI Level (2-5)	MTELP (0-100)	LCT (0-100)	Writing Sample (1-5)
Not accepted	0-32	0-32	0
2	33-44	33-44	1
3	45-59	45-59	2
4	60-79	60-79	3
5	80-100	80-100	4-5

After the beginning of the semester, if students seem misplaced, there is an opportunity to change levels.

Promotion to the next level generally occurs for any students who have performed reasonably well in the class (usually C- or above). It is fairly uncommon for students to be held back from “graduating”. This system leads to a problem which is familiar to administrators in English programs across the country: the students who directly test into a level tend to have higher proficiency than those who are “promoted” into that level. That is, it takes more than one semester’s worth of English study to achieve the gains in proficiency that would increase a student’s score on the placement test by 15 to 20 points. Placement tests are not repeated for students who are already in the system.

3.1.2 The data

The data in this study is a portion of a larger data set collected by members of the Pittsburgh Science of Learning Center (PSLC) and instructors at the ELI. The study has been approved by the University of Pittsburgh Research Conduct and Compliance Office. Please see Appendix C for a copy of the approval notice.

The PSLC data contains samples of written and oral production from classroom activities of students in the English program, as well as standardized test scores. The written data primarily consists of essays assigned as homework in writing and grammar classes; the oral data consists of recordings made during graded classroom activities in speaking and grammar classes.

This study uses Recorded Speaking Activities (RSAs), an oral assessment tool used for speaking classes. RSAs are administered two to four times per semester in every level of

Speaking classes. On the day of an RSA, class is held in a language technology lab so that each student sits at a computer. The students' teacher reads a question on a familiar topic; then the students have two minutes to answer the question and discuss the topic without notes or references of any kind. Students are instructed to speak for the full two minutes. The students record their speech on a Sony HS 90-B microphone headset, using an interface built into the software Runtime Revolution. The sound files are then transferred to a server by means of a file transfer protocol (FTP). When the two minutes have passed, the teacher says "Stop" and the computer automatically stops recording.

The next task is for students to listen to their recording and transcribe their speech. Students are instructed to transcribe every detail of speech, including fillers such as "ah" or "um", self-corrections, and errors. They then have the opportunity to scan their transcript for errors, which they report briefly in a second recording. Finally, they have another chance to record their answer to the given question, keeping the basic text roughly the same but repairing as many errors as possible. The students are graded on both their original recording and their ability to analyze their errors and correct them in the second recording. Content, fluency, and morphosyntactic accuracy are taken into account for assessment; however, it is morphosyntax that usually receives the most attention.

It is the students' transcripts of the first recording—uncorrected—which are used in this study. They have been corrected as necessary to correspond with the sound file exactly.

Topics for each semester are created by supervisors in the Institute. The questions are designed to be appropriate for students of any age, class level, and cultural background; they are also broad enough so that students will not find it difficult to speak for the full two minutes. At

least one RSA per semester is intended to elicit past tense; otherwise, the questions are not explicitly designed to elicit particular structures. A sample question is given in (31):

- (31) Describe something that you liked to do when you were in your country but that you can't do here. Where did you do this? Why did you like it? How did it make you feel?

A full list of the RSA questions used in this study is given in Appendix B. Note that all class levels answer the same questions each semester.¹² No student in the study was given a particular question more than once.

Students have some opportunity to practice their answers to the RSA questions in advance. Approximately a week prior to the RSA, teachers for the speaking classes receive three topics, any of which could be used for the RSA. Teachers are requested to use the topics as ten-minute “warm-up” activities in which students discuss the topics with partners. Teachers generally provide feedback about learner responses, and they may also provide appropriate vocabulary when necessary. However, neither the teachers nor the students know which topic is the “real” one until the day of the RSA.

There are at least two samples of data per semester, generally three or four. For each student, the samples from each semester have been collapsed into a single large group. Combining several samples of data serves several purposes. First, it simplifies the analysis of data by making comparisons between learners of different levels possible. Second, it ensures that each learner will have a substantial amount of data for each semester.

¹² All students receive the same question, with the exception of two RSAs: the first RSA for the first two semesters of the study. These semesters were during the initial launch of the PSLC project, when the study was still in its initial stages.

A third important purpose of combining the data samples is to mitigate the variations in contexts that occur with differing topic prompts. Each semester contains multiple topics, each of which may tend to elicit certain forms or structures. For example, a topic that requires a description of the learner's best friend may elicit third person singular *-s*, while a topic requiring a description of a place may tend to elicit existentials such as *there is*. Collapsing the data from three or four of these questions reduces the chance of the task type influencing the outcome of the analysis. A sample of student data is given in Appendix B.

3.1.3 Participants

Data from 48 of the participants in the PSLC study have been analyzed. Participants were selected from four first language groups: Arabic, Korean, Chinese and Spanish. These languages were selected for two reasons: first, these were the largest groups in the school, and second, the languages have diverse morphosyntactic systems. (Recall that OG and RP predict little L1 transfer.) An effort was made to have a balanced number of students from each first language group; however, it was not always possible to do so, due to the distribution of students in the ELI. An additional goal was to include a balanced number of students from each of the four levels; this proved to be difficult, as very few learners test directly into level 5, and only one learner from the selected L1 backgrounds (a Spanish speaker) did. So that there would at least be a few representative samples from level 5, I included a Thai speaker and a Japanese speaker who tested into level 5. Table 6 below shows the distribution of participants by first language and by level that they first tested into.

Table 6. Participants with L1 and entry level

	Level 2	Level 3	Level 4	Level 5	Total
Arabic	4	6	4	--	14
Korean	5	5	4	--	14
Chinese	2	5	4	--	11
Spanish	2	2	2	1	7
Japanese	--	--	--	1	1
Thai	--	--	--	1	1
Total	13	18	14	3	48

Upon entering the program at the ELI, students filled out a questionnaire with biographical information, including their age, sex, time spent learning English, years spent living in an English-speaking environment, and other languages studied. That information is provided below, in Table 7. When reading the table, note that each student has a number (to preserve anonymity). The letter that precedes each student's number indicates the student's first language (A = Arabic, K = Korean, C = Chinese, S = Spanish, J = Japanese, T = Thai). The final number indicates the level into which the student was initially placed. For example, A45-2 indicates that student #45 is an Arabic speaker who tested into level 2 when he arrived at the ELI.

Table 7. Biographical information of participants

Student	L1	Entry level	Age	Sex	Amount of Eng. Study	Years living in Eng-spkg. envt.	Other langs. known/studied
A45-2	Arab.	2	18	M	5+ years	< 1 year	--
A12-2	Arab.	2	18	M	5+ years	< 1 year	--
A148-2	Arab.	2	19	M	5+ years	< 1 year	--
A25-2	Arab.	2	25	M	5+ years	< 1 year	--
K123-2	Kor.	2	27	F	5+ years	< 1 year	Italian
K156-2	Kor.	2	20	F	5+ years	< 1 year	Chinese
K269-2	Kor.	2	22	F	3-5 years	< 1 year	--
K267-2	Kor.	2	27	M	1-2 years	< 1 year	--
K266-2	Kor.	2	19	F	3-5 years	< 1 year	Japanese
C271-2	Chin.	2	25	F	5+ years	< 1 year	Italian
C126-2	Chin.	2	23	F	5+ years	< 1 year	--
S366-2	Span.	2	53	M	< 1 year	< 1 year	French
S362-2	Span.	2	30	F	5+ years	< 1 year	--
A157-3	Arab.	3	23	M	< 1 year	< 1 year	--
A159-3	Arab.	3	22	M	3-5 years	1-2 years	--
A160-3	Arab.	3	23	M	1-2 years	< 1 year	--
A161-3	Arab.	3	20	M	1-2 years	< 1 year	--
A163-3	Arab.	3	26	M	3-5 years	3-5 years	--
A170-3	Arab.	3	24	M	5+ years	1-2 years	--
A181-3	Arab.	3	24	M	< 1 year	< 1 year	--
K46-3	Kor.	3	23	M	< 1 year	None	--
K101-3	Kor.	3	21	F	1-2 years	None	--
K111-3	Kor.	3	27	F	3-5 years	None	--
K167-3	Kor.	3	27	F	5+ years	3-5 years	--
K300-3	Kor.	3	20	F	1-2 years	1-2 years	--
C282-3	Chin.	3	31	M	3-5 years	< 1 year	--
C177-3	Chin.	3	22	F	5+ years	None	--
C298-3	Chin.	3	26	F	5+ years	< 1 year	--
C301-3	Chin.	3	30	F	5+ years	None	--
C127-3	Chin.	3	30	F	5+ years	< 1 year	--
S173-3	Span.	3	18	F	3-5 years	< 1 year	--
S130-3	Span.	3	32	F	< 1 year	< 1 year	--
A338-4	Arab.	4	30	M	1-2 years	5+ years	--
A279-4	Arab.	4	23	F	< 1 year	< 1 year	French
A199-4	Arab.	4	27	M	5+ years	None	--
A182-4	Arab.	4	27	M	5+ years	None	--
K217-4	Kor.	4	25	M	5+ years	1-2 years	German
K320-4	Kor.	4	29	M	5+ years	None	--
K275-4	Kor.	4	22	M	5+ years	< 1 year	--
K276-4	Kor.	4	22	F	5+ years	None	Chinese, Jap.
C84-4	Chin.	4	28	M	5+ years	< 1 year	--
C118-4	Chin.	4	28	F	5+ years	< 1 year	--
C201-4	Chin.	4	31	F	3-5 years	3-5 years	--
C278-4	Chin.	4	26	F	5+ years	< 1 year	--
S391-4	Span.	4	32	F	3-5 years	1-2 years	Portuguese
S383-4	Span.	4	18	M	3-5 years	< 1 year	German
S100-5	Span.	5	39	M	3-5 years	< 1 year	French
J274-5	Japan.	5	35	F	5+ years	None	French, Korean
T397-5	Thai	5	23	M	5+ years	< 1 year	Chinese

Note that the majority of the students report having had at least five years of English study. Similarly, most report that they have spent less than a year in an English-speaking country; in fact, this generally indicates that they have spent less than a few weeks in the U.S. (i.e., the time between arriving in the country and beginning study in the ELI).

Eight level-5 students, four Arabic speakers and four Chinese speakers, were also interviewed. These students were: A159-3, A45-2, A163-3, A181-3, C282-3, C177-3, C278-4, and C127-3. They were selected for practical reasons; that is, they were the students who were present in the ELI at the time of the study who had been in the program the longest. They were paid for their participation in this project. A description of the interview is presented in the section 3.1.5.

3.1.4 Language background

Both Rapid Profile and Organic Grammar downplay any influence a learner's first language might have on the path of L2 morphosyntactic development, while other theories of second language acquisition (e.g., Full Transfer/Full Access, Schwartz & Sprouse, 1996) predict transfer of first language morphosyntactic elements. In order to assess whether L1 influence is present in L2 development, it is necessary to be aware of the learners' first language morphosyntax. The following sections provide a brief sketch of the languages which are the L1s of the participants in this study: Arabic, Korean, Chinese, and Spanish. The focus is on those aspects of language which are measured by Rapid Profile and Organic Grammar: word order, relative and other embedded clauses, copulas, and tense, aspect, and agreement marking. This information may be useful in explaining certain aspects of L2 development.

3.1.4.1 Arabic

There are many varieties of Arabic, which differ from each other in significant ways. The participants in this study, who are from Saudi Arabia, are speakers of the following dialects: Spoken Gulf Arabic (eastern Saudi Arabia), Hijazi Arabic (Red Sea coast), and Najdi Arabic (southern Saudi Arabia) (Ethnologue, 2007). Additionally, they are all speakers of Modern Standard Arabic, which is spoken throughout most of the Arabic world. Modern Standard Arabic (MSA) is a modernized version of Classical Arabic, the language of the Koran. Since children learn MSA in school, nearly all educated Saudis can understand and speak it. Generally, MSA is spoken in formal situations, such as in news broadcasts or speeches, while the local variety is spoken at home, in the market, and in informal situations (Ethnologue, 2007). For the purposes of this study, it is sufficient to describe the grammar of MSA, as it does not differ from the local varieties in ways that are crucial to this analysis.

Word order. It can be difficult to decide on a canonical word order for Arabic, since the order can be flexible. Both SVO order and VSO order are well represented in spoken and written language. Which order is dominant depends to some extent on the variety of Arabic; for instance, Najdi Arabic is often considered predominantly SVO, while MSA is often considered predominantly VSO (Ethnologue, 2007). Generally, however, the basic or underlying word order of most varieties of Arabic is considered to be VSO, with topicalization accounting for permutations (Suleiman, 1984). Some researchers describe this as a “mixed” VSO/SVO system (Fehri, 1993).

Verbal agreement. In Arabic, verbs must agree in number and gender with the subject; agreement markers are verbal suffixes which fuse number and gender features. For instance, in

the following example with SVO word order, the verb ‘entered’ must be marked as feminine and plural, to agree with the subject ‘women.’

- (31) An-nisaa?-u daxal-na makaatib-a-hunna (Fehri, 1993, p. 34)
the-women-nom entered-f.pl. office.pl.-acc-their.f.
‘The women entered their office.’

An interesting asymmetry in Arabic is demonstrated in VSO order, where the verb must only demonstrate gender marking to agree with the subject; in fact, marking number agreement between the verb and subject is ungrammatical.

- (32) daxal-at an-nisaa?-u makaatib-a-hunna (Fehri, 1993, p. 34)
entered-f the-women-nom officelpl.-acc-their.f.
‘The women entered their office.’

There are two tenses in Arabic: past and non-past (which indicates future and present).¹³ These tenses are marked with an internal vocalic pattern and vowel suffixes, which are added to a consonantal root. For instance, in (33) and (34), vowels are added to the consonantal root *ktb* ‘writing’ to indicate finiteness and tense. (33) shows the verb marked for past tense, while (34) demonstrates non-past; in this case, future is indicated by the adverbial ‘tomorrow.’ Modern Standard Arabic also uses a verbal prefix or modal to express future tense.

- (33) katab-a r-rajul-u r-risaalat-a ?amsi (Fehri, 1993, p. 145)
wrote-3.S.M the-man-NOM the-letter-ACC
‘The man wrote the letter.’

¹³ Another school of thought is to consider Arabic an ‘aspectual language,’ in which case these two forms are referred to as ‘perfective’ and ‘imperfective.’

(34) y-aktub-u r-rajul-u gad-an (Fehri, 1993, p. 145)

3-write-M.S.INDIC the-man-NOM tomorrow-ACC

‘The man writes tomorrow.’

There is a tradition of referring to Arabic as an ‘aspectual language’ because it has no specific aspectual morphology; rather, finite (tensed) forms of the main verb can refer to progressive, perfect or imperfect aspect, especially in embedded clauses (e.g., Cohen, 1989). For instance, consider (35) and (36). In (35), a finite non-past verb is used where English requires a non-finite participial form to indicate simultaneity; in (36), the finite non-past verb ‘play’ indicates ongoing action while the auxiliary verb ‘was’ locates the action in time.

(35) jalas-a y-asrab-u l-xamr-a (Fehri, 1993, p. 147)

sat-3.S.M. 3-drink-S.M.INDIC the-wine-ACC

‘He sat, drinking the wine.’

(36) kaana l-walad-u y-alab-u (Fehri, 1993, p. 148)

was the-boy-NOM 3-play-S.M.INDIC

‘The boy was playing.’ (lit: ‘The boy was he plays.’)

Auxiliary verbs are used to form complex tense/aspect forms. Note that in these constructions there are two finite verbs—a construction that would be ungrammatical in English. For example, in (37), both *yakun* ‘be’ and *hadar* ‘come’ are finite. The tense/aspect is achieved by combining the past tense forms of both verbs.

(37) lam y-akun r-rajul-u (qad) hadar-a (Fehri, 1993, p. 157)

NEG.PAST 3-be the-man-NOM (already) came-3.M.S

‘The man had not already come.’

In (38), both *kaan* ‘be’ and *a?kul* ‘eat’ are finite. In this case, only the auxiliary carries past tense, while the main verb ‘eat’ is non-past.

- (38) *kaan-a l-junuud-u laa y-a?kul-uu-na* (Fehri, 1993, p. 157)
be.PAST-3.S.M. the-soldiers-NOM not 3-eat-M.PL.-INDIC
‘The soldiers were not eating.’

Another way to indicate aspect is through the use of participial forms, which indicate ongoing action in a main clause. An example is given in (39).

- (39) *?anaa musaafir-un* (Fehri, 1993, p. 153)
I traveling-NOM
I am traveling.

Note that in this case no auxiliary verb is required.

Copula. No copula is required in present tense sentences, although a copula is required to indicate past or future.

Modals. Arabic modals occur with finite verbs, unlike in English. For this reason, Fehri (1993) argues that Arabic modals do not belong in INFL; rather, they head a modal phrase (MP) which selects a finite clause. For instance, in (40), the modal verb ‘may’ selects a clause with the finite verb ‘eat’.

- (40) *qad y-a?kul-u*
may 3-eat-M.S.INDIC
‘he may eat.’

Meanings of the modals are similar to English; in fact, replacing the modal ‘may’ with the modal *sawfa* ‘will’ yields future meaning, just as it does in English. Modals may inflect for gender, number or agreement, depending on modal type and position in the clause.

Relative clauses. There are two main types of relativizers in Arabic; the first group, declinable relativizers, must agree in gender and number with the antecedent. For example, in (41), the relative pronoun ‘that’ is marked for masculine gender, singular number, to agree with ‘the boy.’

(41) raʔajtu l-walad-a allathii thaaʔa l-jawm (Galal, 2005, 20)
 saw.1s the-boy-ACC that.3.M.S came.3M.S the-day
 ‘I saw the boy that came today.’

Indeclinable relativizers, e.g., *ma* ‘whatever’, have only one form and are not declined, but otherwise behave similarly.

Note that, as in English, Arabic relative clauses follow the antecedent directly. The relativizer may also be omitted in certain types of clauses, specifically those with an indefinite antecedent, as shown in (42):

(42) raʔajtu walad-an zhaaʔa l-jawm
 saw.1s boy-ACC.(INDEF) came.3s the-day
 ‘I saw a boy who came today.’

Arabic has all the types of relative clauses that English does, including those with prepositional phrases, indirect objects, and genitives (e.g., *whose*), both restrictive and non-restrictive.

Unlike English, Arabic has the option of including a resumptive pronoun in certain types of relative clauses; specifically, definite direct object relative clauses. For instance, in the phrase *the book that you bought*, it is optional to include a resumptive pronoun—literally, *the book that you bought it*.

Other embedded clauses. Arabic noun clauses are similar to those in English. A complementizer such as *?anna* ‘that’ is used to introduce a new finite clause. (43) presents an example.

- (43) 9alim-tu ?anna Zayd-an mariiD-un¹⁴ (Al-Seghayar, 1996, p. 3)
 knew-1s. that Zayd-ACC. ill-NOM.
 ‘I knew that Zayd is ill.’

The noun clause may also appear as a subject complement. In this case, there is no need for the dummy subject ‘it’ as in English:

- (44) az9aja-n-ii ?an ghalab-a Zayd-un Amr-an (Al-Seghayar, 1996, p.3)
 annoyed-1s. that beat-3.S.M.PAST Zayd-NOM. Amr-ACC
 ‘It annoyed me that Zayd beat Amr.’

Adverb clauses, such as the conditional shown in (45), are introduced by a subordinating conjunction and include a finite tensed verb.

- (45) law ?ishtaraytu ?ayyaarat-an jadiidat-an sa-‘abi9u al-qadiimat-a. (Al-S. p. 6)
 if bought-1s. car-ACC. new-ACC. will-sell-I the-old-ACC.
 ‘If I buy a new car I will sell the old.’ (p. 6)

Topicalization. Although VSO is considered canonical word order in most varieties of Arabic, SVO word order appears frequently. This order is thought to be generated through topicalization. In (46) and (47), for example, the DP ‘the children’ is presumably a left-dislocated element which has moved into the specifier position of CP (Fehri, 1993).

¹⁴ [9] is a voiced pharyngeal fricative

(46) al-ʔawlaad-u jaaʔ-uu (Fehri, 1993, p. 27)

the-children-NOM came-3.M.PL.

‘The children, they came.’

(47) al-ʔawlaad-u darab-tu-hum (Fehri, 1993, p. 28)

the-children-NOM beat-I-them

‘The children, I beat them.’

Note that resumptive pronouns remain as verbal suffixes.

3.1.4.2 Korean

The following is a brief outline of relevant morphosyntactic structures in Korean.

Discourse structure. Korean is a null subject language. Subjects may be dropped, as well as any or all nominal arguments, case markers on argument NPs, and even the entire predicate. These omissions occur with “old information,” that is, information that the speaker thinks the listener can recover from context (Kim, 1985).

Word order. Basic Korean word order is SOV, although all six word orders are possible. OSV is especially common, while VSO and VOS are the rarest (Kim, 1985).

Verbal inflection. There is no direct subject-verb agreement of person, number or gender in Korean. However, in certain sentences, verbs carry honorific endings which indicate the relative status of a speaker and a listener.

There are two tenses in Korean, past and non-past. Past tense is expressed through the verbal suffix *-(e/a)ss*, while non-past is unmarked. An example of past tense marking is given in (48); notice the suffix *-(e/a)ss* on the verb *iss* ‘be.’

(48) ku-ka chayk-ul ilk-ko iss-ess-ta (Sohn, 1995, 27)

he-NOM book-ACC read-COMP be-PAST-DEC

‘He was reading a book.’

In (49), the same sentence has a non-past marked verb. Notice the lack of suffix on the verb ‘be.’

(49) ku-ka chayk-ul ilk-ko iss-ø-ta (Sohn, 1995, 27)

he-NOM book-ACC read-COMP be-NONPAST-DEC

‘He is reading a book.’

The suffix *-(e/a)ss* can also mark perfective aspect. For instance, contrast (50) with (51). In (50), the verb ‘bloom’ in the embedded clause is not marked for perfective aspect; the interpretation is therefore that the blooming was not completed when the flowers died.

(50) koch-i phi-taka ci-ess-ta (Sohn, 1995, p. 28)

flowers-NOM bloom-TRANS fade-PAST-DEC

‘The flowers died while they were still blooming.’

However, in (51), the verb ‘bloom’ is marked with the perfective marker *-(e/a)ss*. In this case, the blooming must be interpreted as having been completed by the time the flowers died.

(51) kochi-i phi-ess-taka ci-ess-ta (Sohn, 1995, p. 28)

flowers-NOM bloom-PERF-TRANS fade-PAST-DEC

‘The flowers bloomed and died.’

When used with adverbials of present tense, then *-(e/a)ss* is the equivalent of present perfect in English. An example is given in (52).

(52) Suni-ka cikum mak ttena-ss-ta (Sohn, 1995, p. 28)

Suni-NOM now just leave-PERF-DEC

‘Suni has just left now.’

In those cases where $-(e/a)ss$ is ambiguous between a perfective meaning and a past tense meaning, temporal adverbials such as ‘now’, ‘yesterday’, and ‘tomorrow’ can serve to disambiguate. These adverbials are therefore very important to sentence interpretation.

Using both types of $-(e/a)ss$ together in the form $-(e/a)ss-ess$ indicates past perfect.

(53) Insu-ka mikwuk-ey ka-ss-ess-ta (Song, 2005, p.97)

Insu-NOM U.S.-to go- PERF- PAST-DEC

‘Insu had gone to the United States.’

Notice in (53) that the verb *ka* ‘go’ is marked with $-(e/a)ss$ twice.

Copula. Korean does not use a copula in most contexts where English requires it. For instance, in (54), it is not necessary to use a verb to link the word ‘child’ to the word ‘pretty.’

(54) ku ai-ka yeyppu-ta (Song, 2005, p. 78)

that child-NOM pretty-PLAIN.S

‘The child is pretty.’

Korean also lacks a past tense or future tense copula. Instead, the adjective receives the same inflectional endings as do verbs, indicating time.¹⁵ In (55), for instance, the word ‘pretty’ has both a past tense suffix $-(e/a)ss$ and a homophonous perfective suffix $-(e/a)ss$.

¹⁵ Some researchers consider these adjectives to be stative verbs (e.g., Kim, 2007). This may also be true for Chinese; it is not clear that the categories A, V and P work the same way in Chinese.

(55) Mali-nun han tongan yeppu-ess-ess-ta (Kim, 2007, p. 4)

Mali-TOP one period pretty-PERF-PAST-DEC

‘Mary had been pretty (for a certain period).’

The one place that Korean does use a copula is with predicate nominatives, that is, when a noun phrase attributes a property to another noun phrase in the sentence.

(56) kiho-ka haksayng i-ta (Song, 2005, p. 103)

Keeho-NOM student is-DEC

‘Keeho is a student.’

Note that in (56), ‘student’ is a predicative nominative, also known as a subject complement.

Modals. There are two types of modal expressions in Korean: sentence final and non-sentence-final. Sentence final modality is indicated by the use of sentence endings indicating the speaker’s position in regards to the content of the sentence: surprise, commitment, presumption, reassertion, or insistence. In (57), a sentence final suffix *-keyss* indicates that the statement is based solely on the speaker’s judgment (i.e., epistemic modality).

(57) i chip-un kyewul-ey supki-ka manh-keyss-ta (Wymann, 1996, p. 147)

this house-TOP winter-in moisture-NOM much-JUDGMENT-DEC

‘This house must be very damp in winter’

Non-sentence-final modality is indicated with a complex predicate involving modal markers, aspectual markers, and light verbs, something like the periphrastic modals in English (Jung, 2003). These modals express the equivalent of English *can*, *could*, *may*, *must*, etc. An example is given below.

(58) phyonci-lul nae-sy-o-ya ha-l kos i-p-ni-ta (Wymann, 1996, 127)
 letter-ACC write-HON-csfx-MOD AUX-FUT thing COP-UFS-IND-S

‘You should write a letter.’

ADN = adnominalizer (modifies following noun) AUX = light verb

UFS = upward formality marker IND = indicative

In (58), the suffix *-ya* on the verb *write*, when combined with aspectual marker *-l* on the light verb *ha*, indicates deontic necessity.

Relative clauses. As in most SOV languages, relative clauses in Korean precede the noun that they modify. There are no relative pronouns; instead, Korean uses three types of suffixes which attach to embedded predicates to indicate relativization. These suffixes indicate the time in the embedded clause relative to the time in the matrix clause, that is, *-(u)n* indicates that the time in the relative clause precedes the time in the main clause; *-(u)l* indicates that the time referred to in the main clause precedes the time in the relative clause; and *-nun* indicates that the times are equal. Examples are given below in (59), (60), and (61).

(59) John-i ilk-un chayk-i caymiiss-ta (Kim, 1985, p. 338)
 John-NOM read-REL.PAST book-NOM be.interesting-DEC

‘The book that John read is interesting’

Here, in (59), the marker *-(u)n* means that the reading occurred first; the English translation therefore uses past tense on the verb ‘read’ and present tense on the verb ‘be.’

(60) John-i ilk-nun chayk-i caymiiss-ta (Kim, 1985, p.338)
 John-NOM read-REL.PRES book-NOM be.interesting-DEC

‘The book that John is reading is interesting.’

Here, the marker *-nun* means that the reading is simultaneous with the opinion of the book as interesting; therefore, present progressive is used in both clauses in the English translation. Finally, consider (61) below.

- (61) John-i ilk-ul chayk-i caymiiss-ta (Kim, 1985, p.338)
 John-NOM read-REL.FUT book-NOM be.interesting-DEC
 ‘The book that John will read is interesting.’

In (61) the marker *-(u)l* is used, indicating that the action in the relative clause occurs after the action in the main clause. Here, the reading takes place after the expression of interest.

Restrictive and non-restrictive clauses in Korean have exactly the same form.

Other embedded clauses. Adverb clauses in Korean are formed with a verbal suffix, not with a conjunction as in English. There is a fairly wide range of adverbial endings, for instance, those that express time relations (e.g., *when, while*), reason (e.g., *because*), condition (e.g., *if*), purpose (e.g., *so that*), and others. An example is given in (62), where the suffix *-(e)se* on the verb ‘get rowdy’ has the equivalent meaning to English ‘because.’

- (62) ai-tul-i nemu ttetul-ese cengsin-ul (REF!)
 child-PL-NOM much get.rowdy-because concentration-ACC
 mos chali-keyss-e
 unable obtain-JUDGMENT-INTIMATE
 ‘I can’t concentrate because the children are getting so rowdy.’

In all sentences with adverbial embedded clauses, the embedded clause must precede the main clause.

Topicalization. Sohn (1995) refers to Korean as a “discourse oriented language,” in that discourse contexts play a major role in sentence structure. Korean allows both topicalized

elements and focused elements to move to the left periphery, i.e., the beginning of the sentence. For instance, both of the following are possible:

(63) ponsa-ey-nun yenghi-ka kiho-wa ka-ss-ta (Song, 2005, p. 107)
head.office-to-TOP Yonghee-NOM Keeho-with go-PST-DEC
'To the head office, Yonghee went with Keeho.'

(64) yenghi-ka kiho-wa ponsa-ey ka-ss-ta
Yonghee-NOM Keeho-with head.office-to go-PST-DEC
'Yonghee went to the head office with Keeho.'

In (63) the goal-marked noun 'head office' is topicalized and appears sentence-initially, while in (64) the subject Yonghee comes first and does not need to be marked as a topic. While virtually any nominal element can move to the beginning of the sentence, the verb is nearly always at the end. On the rare occasions that a verb is preposed, it must be heavily stressed and followed by a pause (Sohn, 1995).

3.1.4.3 Chinese (Mandarin)

It is well known that many of the so-called "dialects" of Chinese are in fact mutually unintelligible; for instance, Mandarin and Cantonese differ from each other as much as Portuguese differs from Rumanian, by some accounts (Li & Thompson, 1981). The description of Chinese given here follows the "common language" *Putonghua*, which is taught in schools in Mainland China. *Putonghua* is also taught in Taiwan (where it is known as *guo yu*, or 'national language'), which is the home of many of the Chinese speakers in this study. Note that although these speakers may also use local varieties of Chinese, I will focus on this common variety, which is generally referred to as Mandarin outside China. Mandarin is spoken by approximately 70% of the population of China (Ethnologue, 2007).

The information in this section is taken from Li & Thompson (1981).

Discourse structure. Mandarin is often referred to as a “topic-prominent” language, meaning that the topic—that is, what the sentence or discussion is about—is the first element in the sentence. The topic is assumed to be known to the listener.

(65) zhei ke shu, yezi hen da (Li & Thompson, 1981, p. 15)

this CL tree leaf very big (CL = classifier)

‘This tree, (its) leaves are very big.’

Note that the topic is different than the subject. The topic in (8) is “this tree”, but the subject is ‘leaves.’ Subjects, objects and other elements that appear in the preverbal position are generally interpreted as definite, as they are considered “known” to the listener. For instance if a word such as *ren* ‘person’ appears preverbally, it is interpreted as “the person,” but postverbally it is the equivalent of “some person” (Li & Thompson, 1981).

As in Korean, Mandarin subjects may be dropped entirely if they can be recovered from context by the listener. For instance, in casual conversation it is perfectly acceptable to utter the following:

(66) zuotian nian le liang ge zhongtou de shu (Li & Thompson, 1981, p. 16)

yesterday read PERF two CL hour gen book

‘Yesterday, (I) read for two hours.’

Objects may also be dropped, if they are clear from context.

Word order. Default word order is SVO, but other orders are possible, and SOV is especially common with the so-called ‘ba’ construction in telic sentences. Verb-initial sentences are also possible, where pro-drop permits.

Verbal inflection. There is no subject/verb agreement and no overt tense marking in Mandarin. Rather, time reference is indicated through the use of time adverbials such as “tomorrow”, “right now” or “last year.”

However, there are four aspects in Mandarin: perfective, imperfective, experiential, and delimitative. The perfective, *le*, is used with quantified events, definite events, inherently bounded events, or the first event in a sequence. Quantified events are those that are contained within a particular period of time, for instance “I slept *for three hours*”. Definite events are those in which a direct object is a definite noun phrase, as in (67).

- (67) wo peng - dao - le Lin Hui (Li & Thompson, 1981, p. 192)
I bump – arrive – PERF Lin Hui
'I ran into Lin Hui.'

Verbs with perfective lexical aspect also require the use of *le*; some examples are *si* ‘die’, and *wang* ‘forget’. Sentences with these verbs, such as “She died last year” or “I forgot the address,” are inherently bounded events in that there is an implied end point in each case.

Finally, the first event in a sequence is followed by *le*, as is shown in (68):

- (68) wo chi wan le ni chi (Li & Thompson, 1981, p. 198)
I eat finish PERF you eat
'After I have finished eating, then you eat.'

Notice that the verbs describing the first act, ‘eat’ and ‘finish’, are followed by the perfective marker, while the second use of the verb ‘eat’ is not.

The durative aspect markers are the free morpheme *zai* and the suffix *-zhe*. Durative markers indicate the ongoing nature of an event, similar to the meaning conveyed by the English

verbal suffix *-ing*. For example, in (69), the marker *zai* indicates the continuous nature of the activity of hitting.

- (69) Zhangsan zai da Lisi (Li & Thompson, 1981, p. 218)
Zhangsan DUR hit Lisi
'Zhangsan is hitting Lisi.'

The durative marker may also be used in complex sentences. When *-zhe* appears in the first of two clauses, it signals that the first event is the background for the second. An example is given in (70).

- (70) ta ku zhe pao hui jia qu le (Li & Thompson, 1981, p. 223)
3sg cry DUR run return home go CRS (crs = currently relevant state)
'S/he ran home crying.'

Experiential aspect is marked with the suffix *-guo*. *-Guo* indicates that an event has been experienced, either at an indefinite point in the past, or with respect to a certain reference time. Questions with *-guo* are usually translated into English as "Have you ever...". Statements with *-guo* indicate that something has been experienced in the past, as in (71).

- (71) wo chi - guo Riben fan (Li & Thompson, 1981, p. 226)
I eat EXP Japan food
'I've eaten Japanese food (before).'

Finally, the delimitative aspect indicates that an action has been done "a little bit." It is marked with reduplication of the verb, optionally including the morpheme *yi* 'one.'

- (72) ni shi (yi-) shi kan (Li & Thompson, 1981, p. 232)
you try - (one-) try see
'Try it a little and see.'

The delimitative aspect is often used to make a request more polite.

Copula. The Mandarin copula is *shi* ‘be’. It cannot occur with aspectual markers. Generally, it is used similarly to the English copula, but it also can indicate affirmation; for instance, in (73) the copula can be translated roughly as “it is true that...”

- (73) ta shi mei qian (Li & Thompson, 1981, p. 151)
3sg be not:exist money
‘It’s true that s/he doesn’t have any money.’

Sentences with *shi* are also used where English has existentials such as “there is/are”; the Mandarin equivalent of “There are flowers everywhere” is “Flowers are everywhere.”

Modals. Chinese modals are similar to those in English: *yinggai* ‘should,’ *neng* ‘can’ or ‘may,’ *dei* ‘must,’ *hui* ‘will’, etc.

- (74) ni neng lai ma
you can come Q
‘Can you come?’

Note that modals generally appear with main verbs, although they can appear alone in the case of ellipsis (e.g. *Can you come? I can.*) In (74), the modal *neng* ‘can’ is combined with the verb *lai* ‘come’.

Embedded clauses. The particle *de* appears after a verb to create a noun clause. These noun clauses may serve as complements of nouns, subject complements, or complements of verbs (i.e., direct objects). An example is given in (75).

- (75) ni mei you wo xihuan de (Li & Thompson, 1981, p. 576)
you not exist I like nom
‘You don’t have what I like.’

In this case, the clause ‘I like’ serves as the complement to the verb ‘have.’

The marker *de* is also used in relative clause constructions. In Mandarin, relative clauses precede the noun that they modify. An example is given in (76).

- (76) zhong shuiguo de nongren (Li & Thompson, 1981, p. 580)
grow fruit NOM farmer
‘farmer(s) who grow fruit’

The clause in (76) is a subject relative clause; notice that there is no subject in the clause itself.

Object relative clauses are formed in a similar way.

- (77) tamen zhong de shuiguo (Li & Thompson, 1981, p. 580)
they grow NOM fruit
‘the fruit that they grow’

In (77), the relative clause is “they grow;” and the object “fruit” is “missing” from the relative clause but present in the main clause.

Mandarin also allows relative clauses for instruments and locations (i.e., prepositional phrase relative clauses), time clauses (e.g., “the time when the sun shines”) and reason clauses (e.g., “the reason I bought it”).

Adverb clauses may come either before or after the main clause; the position is dependent on the type of conjunction. For instance, *jiaru* ‘if’ requires the embedded clause to appear first.

- (78) jiaru xia yu women jiu zai wuli chi fan (Li & Thompson, 1981, p. 632)
if descend rain we then at indoors eat food
‘If it rains, we’ll eat indoors.’

The conjunction *jiaru* appears at the beginning of the clause or after the subject or topic, but other conjunctions of this type, such as *de hua* ‘if’, *yihou* ‘after’, and *yiqian* ‘before’ appear at

the end of the embedded clause. Note that *jiaru* also requires the use of *jiu*; thus the two form a correlative pair, similar to English *not only...but also*.

A relationship between clauses can also be established without the use of a linking element.

(79) wo shuo keyi jiu keyi (Li & Thompson, 1981, p. 642)

I say can then can

‘If I say it’s okay, then it’s okay.’

In cases like this, the listener infers the relationship between the two clauses based on context and knowledge of the world. For this reason, sometimes there may be ambiguity or vagueness in the meaning of a clausal relationship.

3.1.4.4 Spanish

Although all of the Spanish speakers in this study are from South America, their home countries vary, and therefore they speak different varieties of Spanish. While there are some morphosyntactic differences between regional varieties, there are no regional differences that significantly deviate from the material presented here.

Word order. Spanish is often classified as an SVO language, and indeed many sentences have this order; however, verb initial structures are also frequent, particularly in the case of unergatives and unaccusatives, as is shown in (80).

(80). Llegó mi nieto. (Hertel, 2003, p. 274)

Arrived my grandson.

‘My grandson arrived.’

Word order is also influenced by discourse factors, in order to place new discourse elements into focus. For example, the sentence in (81) is a natural response to the question “What happened while I was gone?”

- (81) Llamó tu hermana. (Rodriguez, p.c.¹⁶)
Called your sister
‘Your sister called.’

Verb forms. Spanish is mainly a fusional language which uses suffixes to indicate person and number agreement with the subject, as well as tense and aspect. Subject agreement is demonstrated in (82) through (84) below. Notice that person and number marking are fused onto a single suffix which attaches to the root *habl-* ‘speak’.

- (82) (Yo) habl-o inglés.
(I) speak-1ST SING. English.
- (83) (Nosotros) habl-amos inglés.
(We) speak-1ST PL. English.
- (84) (Ella) habl-a inglés.
(She) speaks-3RD SING. English.

A different set of person and number markers are used for different tenses and aspects. The above examples show the form for present tense with simple aspect. Simple past is formed similarly, with suffixes on the verb root. An example is given in (85), where the past tense suffix – *ó* attaches to the verb root *llam-* ‘call’.

¹⁶ Guillermo Rodriguez, University of Pittsburgh

- (85) (Ella) llam-ó ayer
 (She) call-PAST.3RDSING yesterday
 ‘She called yesterday.’

Simple future may also be indicated with a verbal suffix; however, it is common to use the periphrastic construction *ir* ‘go’ + *a* ‘to’ + infinitive to indicate future meaning. An example is given in (86).

- (86) Maria va a llam-ar a Pedro.
 Maria go-3RD.SING to call-INF to Pedro.
 ‘Maria is going to call Pedro.’

Note that this structure is very similar to the ‘be going to’ construction in English. Both constructions can also refer to the so-called “future in the past” (i.e., ‘was going to’) by changing the tense marking of the verb ‘to be’.

There are two aspects that are marked with verbal suffixes: perfect and imperfect. Perfective aspect marks an event that is contained within a certain period of time, while imperfect marks an event that is unbounded by time limits. Thus imperfect is used to describe repeated or habitual actions in the past, as well as events or states that occurred for an extended period of time. An example is given in (87).

- (87) Cuando era niña, (yo) corría frecuentemente.
 When was child, (I) ran.1ST.SING.IMP frequently.
 ‘When I was a child, I ran frequently.’

Perfect aspect is used when there is a clear finishing point to the action (i.e., it is bounded by time.)

(88) (Yo) fui al supermercado ayer.

(I) went-1ST.SING.PERF. to.the supermarket yesterday.

‘I went to the supermarket yesterday.’

The distinction between perfect and imperfect in Spanish does not directly correspond to the distinction between progressive and perfect in English. However, there are compound tenses in Spanish which more closely resemble the corresponding forms in English. For instance, in Spanish the present, past, and future perfect forms are constructed with the auxiliary verb *haber* and a participle, and follow similar usage patterns to English. An example of present perfect is given in (89).

(89) (Yo) he viajado a México muchas veces.

(I) have-1ST.SING.PRES. traveled to Mexico many times.

As in English, the perfect constructions are used to express events that occurred previously at an indefinite time; thus, they cannot be used with specific dates and times.

Another complex verb form is the progressive, which is formed with the verb *estar* ‘to be’ plus a present participle. *Estar* may be inflected for present, past, or future tense to form present progressive, past progressive, and future progressive. The progressive forms are used to indicate that an action is in process, especially if another event interrupts it. An example is given in (90).

(90) Cuando son-ó el teléfono, me esta-ba duchando.

When ring-3RD.SING.PAST the phone 1ST.SING.DAT be-1ST.SING.PAST shower.PART.

‘When the telephone rang, I was taking a shower.’

Note that there are numerous differences as to how progressive forms are used in English and in Spanish. They are far less common in Spanish, and cannot be used to indicate future plans, as in English (e.g., I'm going out tonight).

Copula. Spanish has two copular verbs, *ser* and *estar*. Loosely, *ser* is used with permanent or inherent characteristics (i.e., individual-level predicates), while *estar* is used with changeable characteristics and locations (i.e., stage-level predicates). For example, *estar* is used in sentence (91), because tiredness is not a permanent characteristic of the speaker, but *ser* is used in sentence (92), because it is an unchangeable fact that the speaker is from America.

(91) (Yo) estoy cansada.

(I) am tired.

(92) (Yo) soy Americana.

(I) am American.

Aside from the semantic contrast between *ser* and *estar*, the Spanish copula is used very similarly as the English copula.

Modals. Spanish modality is expressed through full verbs, which (unlike in English) are conjugated for person, number, and tense. Examples for the modal verb *poder* 'can/could' are given in (93) and (94); note that in (93) the verb has a first person singular ending, while in (94) it has a second person singular ending.

(93) (Yo) pued-o escribir libros.

(I) can-1ST PERS.SING write books

(94) (Tu) pued-es escribir libros

(You) can-2ND PERS.SING write books

There has been some controversy in determining which Spanish verbs should be considered modals. Some researchers claim that there are only three true modals: *soler* ‘to be in the habit of’, *deber* ‘must’ and *poder* ‘can/could’ (e.g., Alcina & Blecua, 1975). Others include a large number of verbs, including *quedar* ‘to want’, *saber* ‘to know’, and many more (e.g., Gaya, 1961). These verbs are grammatically similar in that they are followed by the infinitive form of a verb, but researchers disagree whether they actually convey modality in the traditional sense, that is, a speaker’s position in relation to a proposition in terms of its believability, obligatoriness, reality, or desirability.

Relative clauses (RCs). Spanish relative clauses are similar to those in English, in that they are introduced by a relative pronoun, have the same word order as main clauses, and follow their antecedent. All of the relative clause types allowed in English are also possible in Spanish, including possessive RCs and object-of-preposition RCs. In (95) an example is given of a relative clause using *que*, the most common relative pronoun in Spanish.

(95) Pedro tiene los gatos que Maria vio ayer.

Pedro has the cats that Maria saw yesterday

An example of a possessive relative clause using *donde* ‘where’ is given in (96):

(96) Es la parque donde jueg-an los niños.

Is the park where play-3RD SING.PRES the children

‘That’s the park where the children play.’

Relative pronouns can never be omitted in Spanish relative clauses.

Other embedded clauses. Spanish embedded clauses are similar to those in English. For instance, complement clauses are introduced by verbs of cognition. Note that, as in English,

the complement clauses are introduced by a complementizer (*que*), but the structure of the embedded clause is identical to that of main clauses.

- (97) (Yo) sé que Pedro llam-ó anoche
(I) know that Pedro call-PAST.3RD SING last.night
'I know that Pedro called last night.'

Verbs indicating uncertainty or doubt require the subjunctive form of the verb in the embedded clause. Some verbs require subjunctive when used with negative forms, as in (98):

- (98) Maria no cre-e que Pedro est-e aquí.
Maria no believe-PRES.3RD SING that Pedro is-SUBJ.3RD. here
Maria doesn't believe that Pedro is here.'

The complementizer *que* must always be used with finite embedded clauses.

As in English, it is also possible to have non-finite verbal complements. The sentence in (99) demonstrates an infinitival complement to the verb *creer* 'think/believe.'

- (99) Estas mujeres cre-en saber más que nosotros.
These women believe-PRES.3RD PL know-INF. more than us.
'These women believe that they know more than us'
or more literally 'These women believe to know more than us'

Existentials. Spanish uses the verb *haber* 'exist' to form existential constructions. *Haber* conjugates for tense and aspect, but not person or number. For instance, in (100) and (101) it is shown that *hay* is used in present-tense constructions for both singular and plural, masculine and feminine noun phrases. (102) demonstrates that *había*, a form of *haber*, is used to indicate past tense/imperfect aspect.

- (100) Hay dos hombres en la escuela.
There-are two men in the school

(101) Hay una mujer en la escuela.

There-is a woman in the school.

(102) Había una mujer en la escuela.

There-was a woman in the school.

Other forms of *haber* are used to indicate future tense, perfective aspect, or conditionality.

Note that an alternative to *hay* constructions is possible with a number of Spanish verbs, such as *quedar* ‘remain’, *venir* ‘come’ and *ser* ‘be.’ An example is given in (103).

(103) No queda leche.

Not remain milk.

‘There’s no milk.’

It would also be grammatical to form an existential with *hay* with essentially the same meaning as (103), i.e. “No hay más leche.”

Spanish as pro-drop language. An important difference between English and Spanish is that Spanish is a pro-drop language, meaning that subjects are not required in certain discourse contexts (i.e., those in which the subject is clear from context). It is often claimed that pro-drop languages tend to have rich verbal morphology (as does Spanish). Consider (i) below:

(104) Vamos a la playa.

Go-1st plur. pres. to the beach.

We are going to the beach.

Note that the subject (*we*) is not required in Spanish, but it is required in English.

3.1.4.5 Summary of language descriptions

The descriptions of Arabic, Korean, Chinese and Spanish presented above are of course far from being comprehensive. It is not possible at this point to predict all possible influences

speakers' first languages may have on their acquisition of morphosyntactic features. However, the information presented in section 3.1.4 will serve to inform the analysis of morphosyntactic acquisition of English by speakers of a variety of first languages.

The following section describes the measures that were used to examine the acquisition of morphosyntactic features.

3.1.5 Measures

The primary measure in this study is a count of specific morphosyntactic elements in learners' spontaneous oral production. The morphosyntactic elements chosen for the study are those which are predicted by Vainikka & Young-Scholten (1994, 1996a,b, 1998) and Young-Scholten, Ijuin, & Vainikka (2005) to emerge with the building of tree structure (first Inflectional Phrase, then Complementizer Phrase), as well as the features predicted by Pienemann (2003) to emerge with processing development.

Standard principles for analyzing production data were followed. First, any utterances that repeat all or part of the topic prompt are excluded from analysis. For instance, consider the topic given in (105):

(105) Describe **something that you liked to do** when you were in your country but that you can't do here.

To answer this question, one learner described various activities with her friends, and then went on to say the following:

(106) Another **thing that I like to do** there was going to the country. (S383-4, RSA 2)

The relative clause in (106) was not included in this learner's analysis, because she may have repeated it (with a change of pronoun) from the topic prompt.

Another principle that I followed while analyzing the data was to use only the last form or structure in a series. That is, if a student self-corrected a form or structure, then the last one is used, regardless of whether it is correct. For instance, consider the hypothetical utterance in (107).

(107) Yesterday I **fall, falled, fell** down the stairs.

In this case, the speaker is considered to have one appropriate use of past tense. The first two (incorrect) verb forms are not included in the speaker's count.

Both Organic Grammar and Processability Theory rely on the *emergence* of forms and structures, not the *accuracy* of their use. A form or structure is generally considered to have emerged at its first productive use. However, determining which forms and structures are productive and which are formulaic or chunked can be difficult, especially with short data samples such as the ones in this study. Pienemann, Johnston, & Brindley (1988, p. 235) advise eliminating from analysis any forms or structures that demonstrate either of two characteristics. One characteristic is the use of a particular form or structure with only a single lexical item (or pair of lexical items)—for instance, if a learner uses plural forms only with the word *years*, leaving all other nouns in the bare singular form. The second characteristic to watch for is the use of a structure in inappropriate contexts. For instance, a learner might use *don't know* to mean *can't*, *don't understand*, etc. While analyzing the data, any uses like these were eliminated. Additionally, any idioms, common sayings, and other invariant forms were excluded from analysis. Examples are given below:

(108) Examples of phrases excluded from analysis

How are you?

What does X mean?

Let the cat out of the bag

It is of course not always possible to discern whether a learner's use of a particular structure is formulaic, particularly if there is only one example of that structure in the data. For this reason, any time a learner's result relies on a single token (e.g., a single modal verb, a single plural form, etc.), that result is shown as an asterisk between dashes, like this: *-*-*. When there are at least two (different) tokens to base the analysis on, a plain asterisk is shown, and when a large number of widely varying tokens are present, two asterisks are used, like this: ****. This system is used to provide some additional information about usage beyond a simple *yes/no* measure.

While neither Organic Grammar nor Processability Theory makes use of accuracy measures, Organic Grammar does refer to suppliance in obligatory contexts (SOC) in some cases, such as with the use agreement and pronouns. Therefore, the current study includes both measures of emergence and percentages of accuracy. The details of the analysis for each morphosyntactic element are presented in the next section. For each, I have included an explanation of how that particular structure or marking was assessed, and provided examples from the RSA data used in the study.

3.1.5.1 Examining Organic Grammar: Features of VP, IP and CP

The Minimal Trees hypothesis claims that learners transfer a bare VP and its headedness from their first language; then they acquire IP and CP “from scratch” when they notice lexical cues such as complementizers. The following list details the features that are predicted to appear in the initial VP stage, in the IP stage, and in the CP stage. Some of these predictions are taken from the Organic Grammar table, while others are based on Minimal Trees theory directly.

The VP

Word order. Word order of major sentence elements is an important part of the predictions of Organic Grammar, so each learner's word order was examined. Sentences were required to have at least a subject and a verb to be considered; sentences which also included direct objects were especially useful. An example of a sentence with (correct) SVO word order is given below.

(109) She has a son. (C126-3, RSA 2)

An example of a learner using non-SVO word order is given in (110):

(110) Tried my country many kings. (A25-2, RSA 1)

This sentence is classified as having VSO order.

Any deviations from SVO word order were noted in the student's record. It was also noted whether the word order was that of the student's first language.

The IP

Subjects. Since subjects are generally assumed to appear in the Specifier position of IP, they are sometimes taken to be an indicator of the presence of an IP in the grammar. However, Vainikka & Young-Scholten (1994) point out that subjects may be VP-internal for some second language learners, and for that reason are not unequivocal signs of IP. Despite this issue, subjects are considered in this study because Vainikka & Young-Scholten include them in their Minimal Trees theory, predicting that beginning learners will have a bare VP with an optional subject, and that advanced learners will eventually eliminate null subjects.

To be counted as a subject, a noun phrase must appear with a verb in a phrase that can reasonably be interpreted. An example of an acceptable subject is given in (111).

(111) In New Years, **many people** like to play together. (C126-2, RSA 1)

Two productive uses of subjects is enough for a student to be considered to have emerged subjects. However, Vainikka & Young-Scholten (1994) do not use a simple emergence criterion with subjects; rather, they consider an absence of null subjects to be the crucial point that distinguishes stages of development. Therefore, I have measured each learner's suppliance in obligatory contexts for subjects. Learners who have greater than 90% suppliance are considered to have few null subjects, and learners with over 95% suppliance are considered to have none.

Pronouns. Subject pronouns are sometimes taken as a sign of IP because they require IP in order to carry appropriate case marking; for instance, Lardiere (1998a,b) argued that a Chinese speaker named "Patty" had IP in her grammar, based on her perfectly case-marked subject pronouns. Pronouns are also included in the Organic Grammar profile, and are therefore part of this study. Following OG, a pronoun in any position (subject, direct object, indirect object, or object of a position) was considered an acceptable token of pronoun use, although only subject pronouns are theoretically indicative of IP. However, the pronoun was required to appear with a verb in an interpretable phrase, not in isolation. (Note that object pronouns were measured separately as part of an examination of Pienemann's Rapid Profile; see section 3.1.5.2.)

An example of pronoun use is given in (111):

(111) I was about 10 years old (K275-5, RSA 1)

Pronouns that appear as part of stock phrases, such as "How are you" or "I don't know" were disregarded as probable "chunks."

Following the Organic Grammar analysis, pronouns were treated similarly to subjects: only two uses are required to consider them to have “emerged;” however, suppliance in obligatory contexts was measured to indicate whether they were missing. There is naturally some overlap with the measurement of subjects, since pronouns often appear in subject position; however, this measure also includes objects.

Do-support (as in negation). The auxiliary “do” is used in English questions, sentences with a negated verb, and emphasized affirmative statements (e.g., *I do like you!*). Because questions and emphasized affirmatives were rare in this data, do-support counts were generally taken from sentences with a negated verb, as shown in (112) and (113).

(112) I **did not do** anything without his opinion (A181-4, RSA 2)

(113) We **don’t get drunk** by drinking (A338-4, RSA1)

Tense marking on the auxiliary did not need to be correct in order for do-support to be counted. For instance, if (113) had been uttered in a context requiring past tense, it would still be considered to be a case of do-support. However, if the structure of the phrase was significantly altered, particularly because of verb forms, it was considered an “attempt” at marking do-support, and not included in the count. An example of such an attempt is given in (114):

(114) I think we **must don’t** remember independence day. (K123-2, RSA 1)

Attempts such as this might indicate that do-support will emerge soon.

Copula. The copula is a somewhat controversial element in English linguistics. Some researchers suggest that the copula is a normal verb, except that it lacks semantic content or certain syntactic features (e.g., Schütze, 2000); others argue that the copula is the reflex of Infl (the head of IP) when no verb is present (e.g., Becker, 2000). At any rate, any inflected copula

should indicate that IP is present. Young-Scholten, Ijuin & Vainikka (2005) suggest that at first, the only copula form to appear will be *is*, as in the following example:

(115) My favorite holiday **is** Chinese New Year. (C171-2, RSA 3)

As development progresses, other forms of the copula are predicted to emerge: those marked for tense and agreement, as in (116) and (117):

(116) I **was** about 10 years old (K275-5, RSA 1)

(117) My favorite sports **are** swimming and football because **are** very interesting.

(S130-3, RSA 1)

In order to assess the types of copula each learner was using, I made a list of copula forms for each learner. Learners may have no forms, one form, or several. In the results tables, if learners used no copulas or 'is' only, they receive an N. If they use two forms of the copula, they receive an asterisk, and three or more forms of the copula is indicated with a double asterisk. If only one token of a form is present, this is indicated (as usual) with the symbol -*.

Because Organic Grammar does not predict a simple emergence of copula forms, I also calculated the percentage correct, or correct suppliance in obligatory contexts, for each learner. There are a number of possible copula errors. For instance, tense marking could be wrong, as in (118):

(118) One time my best friend said "Okay, M. you and me can I going to the mall. You have a car." So came, my mother led the car and I went driving...I **am** very very scary and at the same time funny. (S273-3, RSA 2)

Agreement marking could also be incorrect, as in (119):

(119) Dogs **is** very famous. (K269-2, RSA 1)

Another possibility is that the copula is omitted entirely, as in (120):

(120) My cat's name Blake. (A264-2 RSA1)

Each learner was given a score on their correct suppliance in obligatory contexts. That is, in order for a learner to achieve a perfect score, she had to supply all required copulas with correct tense and aspect marking. Although each learner's error types were noted while scoring, the suppliance score collapses all of this information into a single measure. Correct suppliance was calculated in standard fashion, by dividing the error-free copula uses by the total number of contexts requiring copula use. The data regarding SOC of copulas was not included directly in the statistical analysis, but rather served to provide background for understanding the data.

Modals. English modals are considered to be in IP, largely because they take no inflectional marking of their own. In this study, each modal (e.g., *can*, *could*, *will*) is counted as one token. In order to be counted, the modal had to appear with a thematic verb in "bare" form, as in (121) and (122).

(121) If baby choose pencil, this baby **will** be smart (K266-3, RSA 3)

(122) We still **can** understand each other (C177-3, RSA 2)

Along with the number of modals, a record was kept of the types of modals a learner used. This measure was necessary to detect whether learners might be "chunking" modal phrases instead of using them productively. The threshold for "emerged" modals was one modal used productively. However, if a learner had only a single use, dashes are used in the table (i.e., -*-) to indicate that it is not certain that the use is productive. The use of at least a modal with different thematic verbs indicates more definitively that the use of modals is productive; therefore, learners with at least two tokens receive a plain asterisk in the table. A learner with multiple modal verbs (three or more) and multiple uses with varying thematic verbs (four or more) is considered to have advanced usage and receives two asterisks.

The modal was considered an “attempt” (and was therefore not included in counts) if it did not appear with a bare verb. For instance, in (123), the meaning is not clear. The modal verb *might* is possibly being used as an adverb instead of a true modal; therefore it is marked as an attempt only:

(123) Might they do, but it is hard is that not the way that we had fun (A160-3, RSA 2)

In (124), the modal verb appears with the auxiliary *do*, and therefore may not be in an IP structure.

(124) I think we **must don't** remember independence day. (K123-2, RSA 1)

The sentence in (124) was also considered an attempt at modal use.

Sentences with expletive (dummy) subjects. Traditionally, *there*-existentials and sentences with *it* as a dummy subject have similar structure. Below, (125b) and (126b) show the simplified underlying structure for (125a) and (126a), based on Stowell, 1978.

(125a) There is a man in the garden

(125b) [e [is [a man [in the garden]]]].

(126a) It is easy to come here.

(126b) [e [is [easy [to come here]]]]

In both of these cases, the subject position is empty (e). Because English requires the subject position to be overtly filled, a “dummy” word with no semantic content is inserted to give form to the empty subject: *it* or *there*.

Note that although expletive subjects are a feature of IP, Organic Grammar predicts that productive use of expletives will emerge in the final stage of acquisition. This anomaly is not explained in Young-Scholten, Ijuin, & Vainikka (2005), but it may have to do with the

assumption that dummy subjects are more “complex” in some way than other aspects of morphosyntax. This issue will be explored in the Discussion, section 3.3.

In this study, *there*-constructions and dummy *it* constructions are counted together. That is, sentences such as the one in (127) are tallied along with sentences such as the one in (128).

(127) **There are** many cities in my country that have poor population (S100-5, RSA2)

(128) **It** was not easy to come here. (K275-4, RSA 2)

As usual, any phrases that appeared to be formulaic were excluded from the analysis.

To qualify as an existential, the structure of the utterance must be relatively intact. That is, it must not include incorrect verbs in the existential structure, nor word order errors involving the existential forms. To give an example, the sentence in (129) does not qualify because the verb *has* is used instead of *is* (or as an alternative analysis, one might consider that *there* is used instead of referential *it*). Therefore, the sentence in (129) is considered an “attempt” to use an expletive.

(129) I like Beijing because **there has** hot spot like Great Wall. (C127-3, RSA 3)

Similarly, an utterance must be interpretable semantically as an existential in order to be tallied as such. For instance, the sentence in (130) is difficult to interpret, as well as having a verbal error.

(130) **There has** many remain in my hand. (C127-3, RSA 3)

This structure was listed only as an attempt at an expletive subject and not included in overall counts.

As with other counts, a single use was indicated in the table with an asterisk between dashes, while several uses were indicated with a plain asterisk. If a learner used both *there* and *it* multiple times, it was indicated with double asterisks.

Tense. Young-Scholten, Ijuin, & Vainikka (2005) use suppliance in obligatory contexts to measure emergence of tense; that is, they do not believe that the simple appearance of tensed forms in students' speech indicates that tense (and therefore IP) has emerged. Rather, they assume that accuracy increases gradually over time.

In this study, learners were given a score based on the suppliance of tense in obligatory contexts. There were therefore two ratings: suppliance of past tense in past tense contexts, and suppliance of present tense in present tense contexts. It was expected that marking in present tense (the assumed default form) would be far superior; for this reason, suppliance of past tense forms was considered to be a more important measure of tense usage.

To measure tense marking, only contexts with finite forms were considered. For example, if a learner used a participial form (such as *-ing*), it was not considered a "context" of use, because it is not clear if the learner intended a progressive form or some other construction. For instance, the sentence in (131) was eliminated as a finite context.

(131) I **shopping** there many times (A45-2, RSA 2)

Obviously, the speech also needed to be interpretable, so that the context for tense was clear. For example, in the following example the learner K320-4 was describing her life as a child. It was therefore clearly a past tense context. In this context the learner uttered (132):

(132) Today, I'm gonna tell my teacher who was very important to me in the past...When I **met** him, he **teach** me how to learn something. (K320-4 RSA2)

For this sentence, the speaker received a score of 1 correct marking out of 2 contexts.

Past tense usage was considered correct even if the form itself was incorrect. For example, the sentence in (133) was considered a correct use of past tense, even though the irregular verb *give* is marked with the regular suffix '-ed'. This reasoning is that it is the

marking of tense, not the lexical form itself, that is of interest in this study. However, the incorrect form was noted.

(133) They **gived** me special task. (A157-3, RSA 2)

For implicational scaling purposes, it was necessary to create a threshold which would divide learners whose tense use has “emerged” from the others. Two thresholds were considered for labeling usage as “emerged”: 60% and 80%. 60% targetlike inflectional use is considered safely above “chance” by Young-Scholten, Ijuin & Vainikka (2005), so it was used here as one measure of emergence. However, research in similar areas generally uses higher percentages, at least 80% suppliance, as an indicator of acquired forms; therefore, 80% was also used as an alternative measure of emergence. In the presentation of data, 60-80% suppliance is indicated with the symbol -*; 80-90% suppliance is indicated with a single asterisk, and 90% marking or over is indicated with double asterisks. Any percentage below 60% is shown with N. At least four tokens of verbs in past tense contexts were required to calculate these percentages. If a learner produced fewer than four contexts, only a slash (/) appears in the table.

Of course, instances of clearly chunked phrases, such as “I come from Taiwan,” are omitted from the analysis.

Aspect marking. Aspect marking is difficult to measure in spontaneous production, because it is often impossible to know what a learner intended by a particular utterance. For example, both (134) and (135) are perfectly grammatical utterances; the difference hinges on the speaker’s attitude towards the duration of the event:

(134) I am living in Chicago.

(135) I live in Chicago.

Without the ability to question the speaker about his intended meaning, it is not possible to judge whether marking is correct or not. Therefore, whenever possible, learners are given the “benefit of the doubt;” i.e., if they use aspect in a reasonable way, I assume that the usage is correct.

To determine whether aspect has “emerged,” I observed each learner’s speech and examined it for any use of aspect marking beyond the “simple” aspects: that is, progressive or perfect. Any use of progressive or perfect was noted, along with whether it appeared to be used appropriately or not. Clear errors, such as the one in (136), were noted as such:

(136) There weren't many kinds of scholarship in my university, but now, **many things change** in my university. (K101-3, RSA 3)

If simple aspect is a default, the error in (136) above is one of omission: either progressive or perfect tense is required.

Another type of error is oversuppliance of aspect marking. For instance, the utterances shown in (137) demonstrate several clear misuses of progressive.

(137) My grandfather of my father is an important person for me because he **was taking** care of us, of all my family in his era and he had a job in the fire station, and he **was earning** a high salary which **is making** a lot of money to build our house (A45-4, RSA 2)

Errors such as these are noted, but are not included in the count of learner’s productive aspect.

If a learner used a present or past participle without an auxiliary verb (*be* or *have*), the verb was not counted as having aspect. The reason behind this choice is that many learners use nonfinite verbs in finite contexts, and it is simply not possible to determine if a particular use is intended as aspectual or not; for instance, consider (138):

(138) I came late. They give me the boarding pass; they **trying** to help me.

(A182-4, RSA 2)

For this reason, the verb had to have a clear auxiliary and a clear participial form to be tallied.

There are contexts in which several aspects are possible; if this occurred, the learner was given the “benefit of the doubt” and the usage was recorded as correct. For instance, the verb *hang* (i.e., *hang out*) in example (t) below could acceptably appear in either simple aspect or progressive aspect.

(139) Pets are not desirable for children. Children should **be hanging** (c.f. **hang**) with their friends. (K320-4, RSA 1)

The utterance in (139) was counted as a use of progressive aspect.

Aspect marking was considered to have emerged when there was at least one appropriate use of progressive or perfect that did not appear to be chunked and was not taken from the topic prompt. An example of such a usage is given in (140):

(140) I was tired; I **was sleeping**. Suddenly, I heard somebody someone knocked on my door. (C171-3, RSA 2)

Special note was taken when especially complex perfect/progressive forms were used, as in (141), or when multiple perfect or progressive forms were used in appropriate contexts along with simple forms, as in (142):

(141) I **had been learning** a lot of things from my group. (C201-4, RSA2)

(142) In that time, I **hadn't driven** a car for a long time. He **let** drive the car. We **was driving** in the highway. I **was driving** faster than 120 Kmph. Then we, I **make** the speed slower...suddenly the hood **flew** up... A12-3, RSA 2)

As in other cases, if there was only one use of progressive or perfect marking an asterisk in parentheses was used in the table. Additionally, if none of a learner's uses of aspect were in contexts that clearly and unambiguously required progressive or perfect aspect marking, then an asterisk with parentheses is used to indicate uncertainty about the productivity and appropriateness of the marking. For example, a learner with several utterances similar to the one in (141) would have an asterisk with parentheses. Learners with over three uses of appropriate aspect marking were given two asterisks to indicate advanced use.

Agreement. As with tense, Organic Grammar does not treat single uses of agreement marking as indicators of emergence. Rather, a percentage of suppliance in obligatory contexts is taken to see whether learners' marking is improving, and if they have correct marking at a level above chance.

Agreement in English thematic verbs is limited to third person singular *-s* marking. There are two possible errors with this marking: an overuse of *-s* marking, as in (143), and omitted *-s* marking, as in (144). It was assumed that omitted marking would be more common, as the "bare" form of the verb is generally assumed to be a default form.

(143) I **thinks** shopping in my country have has a different in USA. (A45-3, RSA 1)

(144) My best friend is Z...she **like** to talk about something with me. (C126-3, RSA 1)

In order to calculate percentages of suppliance for each of these contexts, a count was taken of correct uses of marking, and that number was divided by the total number of contexts. In the statistical calculations, only suppliance in contexts requiring 3rd person *-s* were used, because these contexts require a non-default form. At least four tokens of verbs in 3rd person singular contexts were required to be included in the statistical calculations. If a learner produced fewer than four contexts, only a slash (/) appears in the results table.

Another type of agreement error involves the use of non-finite forms in finite positions, with no apparent aspectual meaning. For instance, C171-3 uses an auxiliary plus bare verb in a large number of contexts:

- (145) People return their home with family together. They **are eat** together, they **are play** together, they **are talk** about together, they **are do** a lot of things together. They are happy. (C171-3, RSA 3)

These errors were included in overall suppliance percentages. The speaker in (145) would therefore be considered to have little agreement for plural nouns in this sample.

Sentences with non-initial subject. Organic Grammar includes the number of sentences with non-initial subjects as an additional clue to learners' grammatical development. Sentences which begin with a prepositional phrase or other non-subject element are included because they imply the presence of a position before the subject. This position may be in IP or CP.

For this part of the study, a count was taken of the number of times a learner began a sentence with a prepositional phrase or subordinate clause. An example of such a sentence is given in (146):

- (146) **In 1886**, the first fundamental letter of rights was wroten in my country. (S100-5, RSA3)

Sentences which begin with adverbs such as *actually* or *sometimes* were not included in this count, because they often appeared to be disconnected from the main clause by a pause. Thus they may be adjoined by simple parataxis, or simply used as filler words. This methodology is more conservative than that proposed by Young-Scholten, Ijuin & Vainikka (2005). However, a more liberal count was taken with the Rapid Profile measures and was available for use in statistical calculations. (The more liberal count is described in section 3.1.5.2.)

The CP

We now move on to CP structures, which are predicted by Minimal Trees to appear after IP has emerged.

CP Complements. This is a very diverse category involving a range of structures, all of which require a Complementizer Phrase. One type of CP complement is a noun clause. These are generally complements of verbs, although they can also appear as complements of nouns. An example from the data is given in (147); note that the speaker has omitted the optional complementizer *that*.

(147) I think **this is a good custom for the baby** (C282-4: RSA 3)

This structure typically involves verbs like *know* or *think*, as when presenting an opinion. Because many RSA topics ask students to present their opinions, this was a relatively common structure.

Another type of CP complement is an infinitival complement to a verb. These typically follow verbs such as *want*, *like*, or *try* in English. An example is given in (148).

(148) I **want to see** them. (Korean 275-4, RSA1)

If a learner only used one verb with infinitival complements, then this was considered a possible case of chunking. For instance, one student (C126-2) uses *like* with an infinitive four times in a single RSA recording, where few other matrix verbs are used. (Importantly, however, different verbs appeared in the infinitive phrase.) In this case, the use of a complement clause was noted, but it was marked with dashes (-*-) to indicate that it could be a chunked pattern and not productive.

Note that infinitives may also complement adjectives or nouns. For instance, the sentence in (149) contains an infinitival complement to the adjective ‘happy’.

(149) I’m happy **to help** you.

Although structures like these were less common than complements to verbs, they were included in the analysis. Infinitival complements may also be preceded by a *wh*-word, as in (150).

(150) They want to know **how to sentence** the problem (C282-3, RSA3)

Some learners made errors in the infinitive phrase, as shown in (151):

(151) I **like watch** this sports on TV (S383-3 RSA 1)

Utterances like the one in (151) were included in counts, with the error noted. However, the complement was required to have a verb in order to be counted. For instance, consider (152):

(152) I **want to back** to Taiwan to see him. (C271-2 RSA2)

Because the error involves the absence of a verb form, the utterance in (152) was considered only an “attempt” at an infinitive complement.¹⁷

A third type of CP complement involves what Pienemann refers to as a “Cancel Aux-2nd” structure. These are noun clauses which are introduced by a *wh*-word, similar to a question but with standard (non-question) word order. An example is given in (153):

(153): They [pets] can understand **what we are talking about** (C278-4, RSA 1)

These types of clauses are especially important in Rapid Profile, and are discussed further below.

Despite the diversity of possible complement clauses, they were counted similarly to other measures; that is, one token is shown with an asterisk in parentheses; several uses with a

¹⁷ It is possible that the learner intends a verb here, i.e., “back” = “return” in the learner’s interlanguage. However, to maintain consistency when analyzing results, I have avoided speculating about learner intention.

plain asterisk, and multiple uses (4) of multiple kinds of complement clause are shown with two asterisks.

Adverb (adjunct) clauses. These clauses provide adverbial meanings regarding time, place, condition, etc. Unlike complement CPs, they can generally be omitted with no loss of grammaticality. These clauses are included under the category of “complex subordination” in the Organic Grammar system. Two examples from student data are given below.

(154) **If baby choose pencil**, this baby will be smart. (K266-3, RSA3)

(155) **When I finished my classes**, my mother led me, led her car. (S173-3, RSA2)

In order to consider a learner to have “emerged” adverb clauses, she must have at least one productive use. Learners with two or more uses receive a plain asterisk, while learners with multiple (4) adverb clauses with different conjunctions are marked as having advanced usage.

Because clauses. Generally, clauses with *because* are considered to be adverb clauses, like clauses with *if* or *when*. However, Young-Scholten, Ijuin & Vainikka (2005) claim that clauses with *because* are less complex than other adverb clauses, naming them an aspect of “simple subordination.” They provide the following example:

(156) Someone’s die **because he have accident**. (YS,I, & V, 2005, p.13)

They contrast this *because*-clause with other types of adverb clauses, which they call types of “complex subordination.” An example of this kind of complex subordination is provided in (157):

(157) **When you reverse**, you have to see anybody behind (YS,I, & V, 2005, p.13)

Although it is not explained in YS, I, & V’s article why this division is made, the notion may stem from the intuition that *because* can be used—fairly straightforwardly from a pragmatic viewpoint—as a way to continue a thought or develop an idea. Developing an idea is something

that these learners must do frequently, since they must speak for an uninterrupted two minutes. Consider, for example, the speech from a level 4 Spanish speaker in (158). The speaker begins by explaining that she felt more freedom in her home country than she does now. Then, to elaborate the idea, she uses a *because* clause:

(158) We had more freedom to do things that you want there...**because you live there, you know the place and know the friends and all that.** (S383-4, RSA 2)

In the above example, the *because* clause appears after a slight pause, with a marked change in intonation. It is possible that the word *because* is serving not as a subordinating conjunction, but rather as a transitional adverbial similar to *after all* or *well*.

Of course, *because* does not always serve as an adverbial. For instance, it appears to function as a subordinating conjunction in the example given by Young-Scholten, Ijuin, & Vainikka, repeated below in (159):

(159) Someone's die **because he have accident.** (YS,I, & V, 2005, p.13)

One solution would be to separate uses of *because* that appear to be true subordination from those which are not. However, this may prove to be nearly impossible in many cases. For instance, it would be very difficult to ascertain the "true nature" of the *because* clause in (160):

(160) When I was 11, my elementary school teacher was very important to me, **because she is always nice to me, and she make me change.** (K276-4: RSA 2)

To avoid potentially making errors in classifying adverb clauses, and to be consistent with the techniques used in Organic Grammar, all uses of *because* clauses were counted separately from other types of clauses.

To be counted as an instance of *because* subordination, the conjunction *because* had to appear with a verb, as in (161):

(161) I don't stop ballet **because my teacher said "you can do it!"** (K266-2: RSA 2)

Other errors in the clause, such as this case of an omitted subject, were ignored when evaluating learners' performance on clauses:

(162) I like Saudi Arabia **because is very very nice country** (A181-3: RSA 1)

Utterances with *because* but no verb were not counted. For instance, the utterance given in (163) was not considered a token of a *because* clause.

(163) In the U.S., people mmm, **because**, mmm. (C127-5, RSA 2)

Because-clauses were counted in the same manner as adverb clauses.

Relative clauses. There are four main types of relative clauses in English: subject, object, possessive, and adjunct. In subject relative clauses, the relative pronoun serves as the subject of the subordinate clause. An example is given in (164):

(164) We want to have another president **who can lead us to good economic year, situation.** (Chinese level 5, RSA 3)

In an object relative clause, the relative pronoun serves as the direct object, indirect object, or object of a preposition. An example of a direct object clause is given in (165):

(165) I have to count the steps **which the horse have to take.** (A338-4, RSA 2)

Adjunct relative clauses are those in which the relativized element functions as an adjunct, usually providing information about time, place, or manner.

(166) I'm just starting to make friends and know the place, knowing Oakland and knowing Forest Hills, the place **where I live.** (S383-4, RSA 2)

Possessive relative clauses were not present in the data.

Young-scholten, Ijuin, & Vainikka (2005) consider subject relative clauses to be "simple subordination," while other types of relative clauses are "complex." Therefore, I have

categorized each student's relative clause use by type. As usual, relative clauses that repeat part of the topic or appear to be chunked were not considered.

There were also a few cases where it was not clear whether a relative clause was intended or not. An example is given in (167).

(167) There is a lot of events **happened to me** (A161-3, RSA 2)

In these cases, the structure was listed as an "attempt" and not included in overall counts.

One clear, productive use of a relative clause was enough to consider the learner to have "emerged" relative clauses in her grammar. Learners with two relative clauses receive a plain asterisk, while learners with multiple relative clauses of varying types (e.g., subject, object, etc.) received two asterisks.

Questions. Questions are an important part of both Organic Grammar and Pienemann's Rapid Profile. Unfortunately, they were too infrequent in the data to use as a diagnostic tool, so they were left out of the study. I will return to the issue of including questions in studies such as these in the Discussion (section 4).

Passive. Passive is referred to as an aspect of "complex syntax" in Organic Grammar. This is a reasonable assumption; traditionally, Generative theory assumes that passive morphology suppresses external theta-role assignment and case marking to the object, forcing DP-movement of the object to subject position. Even from an atheoretical viewpoint, passive voice may be more difficult than active voice because it requires the theme (typically in object position in English) to appear in subject position, while the agent (typically the subject in English) is omitted or included in a prepositional phrase. Unfortunately, it is not clear why passive should be considered "complex" in Minimal Trees theory, as it does not require CP; however, it was included in this study so as to be as complete as possible.

Passives were analyzed as follows. First, errors in the form of the participle or auxiliary were not considered to be errors in passive marking. For instance, (168) was scored as a correct use of passive.

(168) In 1886, the first fundamental letter of rights **was wroten** in my country.

(S100-5, RSA 3)

Similarly, if the participle is present but the auxiliary is omitted, the utterance was counted as a token of passive use, but it was marked as having an auxiliary error. An example is given in (169).

(169) I **invited** by US council...I **invited** to United States because I have an international competition in California (A199-4, RSA1)

However, so-called “false passives” with verbs that are not able to be made passive in English (unaccusatives) were not included in overall counts of passive use, although I made a note of their use. An example is given in (170).

(170) He **was died** before three years ago. (A144-2, RSA2)

Passive phrases that are likely to be chunked, especially statives such as *it's called* or *it's located* were not included in the count, although they are noted. In order for a student to be scored as having “emerged” passive voice, the student must have used at least one clearly non-chunked passive form in an intelligible sentence. Learners with only one use, or with auxiliary errors, are given an asterisk with parentheses. Learners with several uses receive two asterisks.

It is important to note that there is a limitation with using passive as a measure with spontaneous speech data, because passive was used relatively infrequently even by advanced learners, and contexts were not always conducive to its use.

Other elements. Other elements not directly related to VP, IP or CP were included in this part of the analysis because they are used in the Organic Grammar table.

Coordination. Coordination in itself is not associated with the presence of IP or CP; rather, it links two similar phrases. Coordination is predicted by OG to appear relatively early in development.

When learners connected two independent clauses with the use of a coordinator such as *and*, *but* or *or*, it was counted as a token of coordination. Compound subjects or predicates were not included in the count. In order to be counted as a coordinator, the connected clauses needed to have verbs and to be intelligible. An example of a token of coordination is given in (171):

(171) I walk to door, **but** I nervous. (C126-3, RSA 2)

Notice that errors in the independent clause do not affect the way that coordination is measured in this study.

Size of sample in AS-units. A count of AS-units (Analysis of Speech units) (Foster, Tonkyn & Wigglesworth, 2000) was conducted on each sample. An AS-Unit is a main clause plus any additional subordinate clauses. Only utterances with verbs are counted, and repetitions and false starts are not included in AS-counts. The main purpose of this count is to have an idea of the size of the sample and to provide a basis for measurement of overall complexity.

Vocabulary measure. Following Young-Scholten, Ijuin, & Vainikka (2005), I counted the total number of unique nouns used and the number of times these nouns were repeated for each sample. These two numbers together provide insight into each learner's level of vocabulary. This information was not included in statistical measures.

3.1.5.2 Testing Pienemann's Rapid Profile

Data to examine Pienemann's Rapid Profile was collected with the same procedure as described in Section 3.1.5.1 above. Certain measures, such as 3rd person singular *-s* marking, are relevant to both RP and OG. Measures that were taken specifically to test Rapid Profile are listed below.

Words, formulae. Any learners who only used individual words or "chunks" would have scores of zero on all counts of syntax, including AS-Units.

Plurals. Learners' use of nouns was observed to discover whether they used plural *-s* marking, as in *dogs*. Since Processability Theory relies on emergence of marking, not percentage correct, learners were scored as either having plurals or not. For this purpose, only regular forms were counted (e.g., *cats* but not *children*). An example of a productive plural form is given in (172):

(172) I like small and cute **dogs**, because they are very kind. (C271-2, RSA 1)

The use of *dogs* would not, of course, be counted as a token of plural if it appeared in the following idiom:

(173) It's raining cats and **dogs**.

Learners with only one example of plural marking are marked with an asterisk in parentheses to indicate that there is some doubt whether they have productive plural. Two or more uses of plural (on different nouns) is indicated with a single asterisk; learners with multiple uses of plurals on multiple lexical items receive two asterisks.

Poss *-s* on nouns. There were not enough examples of possessive markers on nouns in the learner data to use this measure in this study. Presumably, this absence of forms is due to the task (i.e., spontaneous response to certain topics).

Use of –ing. Following Pienemann (1998), any productive verbal use of –ing was considered an acceptable token of –ing use, regardless of whether it appeared to be used correctly. An example of such a token is given in (174):

- (174) My father is an important person for me because he was **taking** care of us
(A45-4, RSA 2)

Nouns such as *shopping* or *building* were not included in these counts, under the assumption that these were likely to be unanalyzed wholes. For example, the word *building* in (175) is not included in any counts.

- (175) There is a lot of **building** there (A45-2, RSA 2)

Uses of –ing were marked in the same way as other elements. That is, if a learner only had one use of –ing, she received an asterisk with parentheses. Two or more uses were marked with a plain asterisk, and multiple uses with multiple lexical items was marked with a double asterisk.

SVO order and -ed marking. Both of these counts were taken from the Organic Grammar measures.

Adverb 1st. This measure is similar to the measure in Organic Grammar regarding sentences with non-initial subjects; however, the approach here was broader. That is, any preposed adverbs or other adverbials were considered to be tokens of Adverb 1st, not just prepositional phrases and subordinate clauses. An example is given in (176):

- (176) **Actually**, my mother, my mom also doesn't like pets. (K217-5, RSA 1)

The adverbial was required to precede an interpretable clause with a verb to be counted.

S neg V(O). Pienemann notes that this structure is common in developing learners who are beginning to use negation. A hypothetical example is given in (177).

- (177) I no like that.

This structure did not occur in any of the learners, possibly because they have all advanced beyond this stage; therefore it is not included in the table.

Poss pronoun. Uses of possessive pronouns such as *my*, *your*, and *their* were noted. Possessive pronouns that occurred in formulaic phrases, such as *my name is...* or *what's your name?* were excluded from analysis. Counts were conducted in the same way as with other forms.

Cancel Aux-2nd. This label refers to complement clauses with *wh*-words. An example of an utterance with *Cancel Aux 2nd* is given in (178).

(178): They [pets] can understand **what we are talking about** (C278-4, RSA 1)

The reason that Pienemann refers to this structure as “Cancel Aux 2nd” is that it requires the *wh*-word to be displaced from its position, as with a question; however, the auxiliary verb does not appear before the subject, as it does in questions. The idea is that learners must “un-learn” the question rule in order to produce this form.

One use of a Cancel Aux-2nd form is indicated with an asterisk in parentheses. With two or more uses, a single asterisk is used. Several occurrences with different *wh*-words are indicated with two asterisks.

V-Particle. This label refers to phrasal verbs such as *pick up*, *look up* or *get over*. Very few of these occurred in the data, and the only phrasal verbs that were used were intransitive (e.g., *grow up*, *get together*). Therefore they were left out of the table.

Question measures. In Rapid Profile, there are many measures that test a learner's ability to form questions. Unfortunately, there were not enough questions in the data to use any of those measures in the implicational table.

Note that the methodology used here, observations of spontaneous production, is not the same as is used with Rapid Profile. Rapid Profile advocates eliciting certain structures, such as questions. The fact that particular structures were not elicited in these data means that only a percentage of Rapid Profile measures can be tested. However, Pienemann (1998) asserts that the emergence principles elaborated in Processability Theory should apply to learner production in all task types. The results of the current study are important because I explore the question of whether Rapid Profile elements can be used in a placement test based on spontaneous production data.

3.1.5.3 Statistical procedures.

The primary statistical procedure used in this study is implicational scaling, as outlined by Hatch & Farhady (1982). All counts of morphosyntactic elements were entered into two Excel tables—one for Organic Grammar and one for Rapid Profile. The elements predicted to appear earliest are on the left, and the elements predicted to appear last are on the right. For instance, the use of correct SVO word order is one of the first predicted milestones for an L2 learner in the Organic Grammar system, so the column with those results is the leftmost one. Complex subordination (e.g., adverb clauses) are predicted to appear last, so that column is on the far right. Each asterisk, double asterisk, or asterisk in dashes (-*-) is counted as one point, and the total points are tallied for each learner. The data are then sorted by number of points per learner, so that the learners with the lowest number of points appear at the top of the table, and those with the highest number of points appear at the bottom.

Learners who have the same point tally may be rearranged so as to reduce the number of errors in the table. That is, it is acceptable to change the order of learners so that the table has the best fit as an implicational table (Hatch & Farhady, 1982). Similarly, the elements that are

predicted to occur in the same stage of learning may be rearranged to as to have the best fit. For instance, in Rapid Profile, *-ed*, *-ing*, and plural *-s* are all predicted to emerge in stage 2. These three elements may be rearranged within that stage so as to achieve the best table.

A line is then drawn on the table to separate the “emerged” side from the “not yet emerged” side. Ideally, this line is a diagonal line from the upper left corner to the bottom right corner. As a point of reference, an ideal implicational table is presented below:

Student	A	B	C	D	E	F	G	Points
1	N	N	N	N	N	N	N	0
2	*	N	N	N	N	N	N	1
3	*	*	N	N	N	N	N	2
4	*	*	*	N	N	N	N	3
5	*	*	*	*	N	N	N	4
6	*	*	*	*	*	N	N	5
7	*	*	*	*	*	*	N	6
8	*	*	*	*	*	*	*	7
Points	7	6	5	4	3	2	1	

Figure 14. An ideal implicational table

An error occurs in the table when there is an “N” on the “emerged” side or an asterisk on the “not emerged” side of the line. Consider the figure below, for instance. Student #5 has an N out of place and an asterisk out of place. Each one of these is an “error” in the table.

Student	A	B	C	D	E	F	G	Points
1	N	N	N	N	N	N	N	0
2	*	N	N	N	N	N	N	1
3	*	*	N	N	N	N	N	2
4	*	*	*	N	N	N	N	3
5	*	N	*	*	N	*	N	4
6	*	*	*	*	*	N	N	5
7	*	*	*	*	*	*	N	6
8	*	*	*	*	*	*	*	7
Points	7	5	5	4	3	3	1	

Figure 15. An implicational table with two errors

There are two statistical procedures that are used to calculate whether a set of data is a good implicative table or not. The first calculation is the Coefficient of Reproducibility (C of R), which indicates how predictable the results are for all the individuals. The C of R is calculated by first adding the number of errors in the table. Then the number of errors is divided by the number of squares in the table, and subtracted from 1. The equation is given in (179), and an example based on Figure 15 is given in (180).

$$(179) \text{ Coefficient of Reproducibility} = 1 - (\# \text{ of errors} / \text{total number of responses})$$

$$(180) \text{ C of R} = 1 - (2/56)$$

$$\text{C of R} = 1 - .036$$

$$\text{C of R} = .96$$

A Coefficient of Reliability over .9 is generally considered to be indicative of a valid implicative table (Hatch & Farhady, 1982).

The second calculation is the Coefficient of Scalability (C of S). There are a few steps to this procedure. First, all of the emerged (i.e., those marked with an asterisk) responses are tallied and divided by the total number of responses. This number is the Minimal Marginal Reproducibility (MMR). The equation is given below in (181), and the calculation is performed for Figure 15 in (182).

$$(181) \text{ MMR} = \text{number of emerged responses} / \text{total number of responses}$$

$$(182) \text{ MMR} = 28/56 = .5$$

The difference between the C of R and the MMR is then computed. This figure is called the Percent Improvement in Reproducibility (PIR). The equation is given in (183) and the example for Figure 15 is computed in (184).

$$(183) \text{ PIR} = \text{C of R} - \text{MMR}$$

$$(184) \text{ PIR} = .96 - .5 = .46$$

Now it is possible to calculate the Coefficient of Scalability. This figure is intended to indicate if a set of variables is truly scalable. It is calculated as shown in (185), and the example is continued in (186).

$$(185) \text{ C of S} = \text{PIR} / (1 - \text{MMR})$$

$$(186) \text{ C of S} = .46 / (1 - .5)$$

$$\text{C of S} = .92$$

The coefficient should be over .6 or .65 for the table to be considered scalable.

The Coefficient of Reproducibility can be calculated for individual items, as well. This figure provides information about how predictable the emergence of a single element is. The formula is given in (187):

$$(187) \text{ C of R per item} = 1 - (\# \text{ of errors per item} / \# \text{ of students})$$

For instance, for item B in Figure 15 above, the C of R is calculated as follows:

$$(188) \text{ C of R per item} = 1 - (1/8) = .88$$

The C of R should be over .9 for the item to be considered predictable. The reason that one error caused there to be a lower C of R in this case is that the number of students (8) is very low.

Hatch & Farhady (1982) point out that there are many problems with using implicational scaling. First, it requires researchers to decide on a “cutoff point” for suppliance in obligatory contexts for certain features. For instance, a researcher might decide that a learner who uses plural *-s* 60% of the time has acquired plural, or the researcher might decide on a much higher cutoff point, such as 90%. Changing the cutoff point may strongly affect the results. In order to test the claims of OG, I use the cutoff points given by the authors of OG: 60% accuracy.

Another problem occurs with naturalistic data such as those used in this study. In naturalistic data, there may not be context for certain features. For instance, a learner may not happen to use any plurals during a given recording. If data is missing, the researcher must decide what to do with those squares in the table. In this study, I have chosen to treat them as absent from the table. That is, they are simply not included in any counts, either as an error, or as a correct response, or in the total count of responses. However, because a missing data point can interfere with a learner's point count, I have allowed more flexibility in changing rank order when data is missing. For instance, a learner who has a total point count of 4 but who has two squares of missing data may be moved into the group of learners with counts of 5 or 6 (but not 2 or 3, or 7 or 8), because if more data were available, the learner could have appeared in either of those groups. This adjustment ensures that the best possible fit for the table is ensured, while the integrity of the results remain uncompromised.

Hatch & Farhady point out a final issue, regarding cases where there are only a few tokens of data. For instance, if a learner only has two contexts for the plural, and one of them is marked incorrectly (i.e., the singular form is used), should the researcher eliminate the data completely, or treat it as 50% suppliance? For this study, I have chosen four tokens of a particular form as a minimum number of occurrences when calculating suppliance in obligatory contexts. For instance, if a learner only uses three past tense verbs, those data are eliminated, as if the learner had not had any context for past tense at all. This practice is intended to ensure that a reasonable number of tokens is used to calculate percentage of suppliance. Note that four occurrences are only required when dealing with percent measures; all other measures are based on emergence, where one productive form is enough to consider the element acquired.

Once calculations were made for Organic Grammar, the elements were rearranged to find the best possible implicational table with these data. This was accomplished by adding the total points for each element and then sorting the data so that the lowest number (i.e., the earliest emerged element) was on the leftmost side of the table and the highest number (i.e., the last emerged element) was on the rightmost side of the table. The Coefficient of Reproducibility and the Coefficient of Scalability were calculated for this new table.

3.1.5.4 Interviews

In addition to the data described above, I collected interview data from eight students in the intensive English program. The purpose of these interviews was to provide a more in-depth, qualitative view of the individual factors which could influence learners' progress: the learners' motivation, time spent practicing English and interacting with English speakers, attitudes to the United States and other English-speaking countries, and views about language learning and grammar. This information then serves to inform an analysis of the learners' morphosyntax, using Meisel, Clahsen & Pienemann's (1981) and Clahsen, Meisel & Pienemann's Multidimensional Model.

The eight students were selected based on their willingness to participate in the study, the length of their time in the English Language Institute (at least three semesters), and their first language: Chinese, Korean, Arabic or Spanish. They were also required to be in level 5, the highest level in the institute, so that progress over time could be observed. Students were solicited with a personal letter that arrived in their grammar class and explained the interview procedure and the monetary compensation (\$20). No Korean or Spanish speakers were available for the study, but four Chinese speakers and four Arabic speakers agreed to participate.

The interviews lasted between 25 minutes and 50 minutes, depending on the conversational style of the interviewee. The format of the interview was relaxed and informal, so as to encourage students to be open about their experiences and attitudes. Although each student was presented with the same scripted interview questions, I also posed unscripted follow-up questions when I thought that elaboration of a certain point was required, or when it seemed that exploration of a certain area would provide further insight into the student's experiences. The questions used in the interview are presented in Appendix C.

The interviews were transcribed and examined for three main factors: motivation to learn and effort in the ELI; amount of interaction with English speakers, practice in English (listening, speaking, and reading); and attitude to language learning and grammar. Two learners (one Chinese speaker and one Arabic speaker) with high motivation, positive attitude, and the most interaction with native speakers were selected for further study, and two learners with lower motivation, negative attitude, and the least interaction with native speakers were also selected. For convenience, I will refer to the two learners with higher motivation and more interaction as "integrative" and the two learners with lower motivation and less interaction as "less integrative." These are the terms used by Meisel, Clahsen & Pienemann to refer to the two types of learners in their Multidimensional Model. However, it should be understood that I have used different methods and criteria to define the two groups.

First, I present a description of the two groups of learners, with information taken from the interviews. The integrative learners were C127-3, a Taiwanese student, and A181-3, a student from Saudi Arabia. C127's reason for learning English was to obtain a better job in Taiwan than her current one. While studying in the U.S., she worked in a Chinese restaurant and enjoyed interacting with the customers, who were nearly all English speakers. She had an

American conversation partner (through an ELI program) who she met with two hours a week, and she participated in many ELI activities where English was spoken. She also enjoyed socializing with ELI students outside of class, and always tried to speak English with them. She read magazines and newspapers for pleasure, and also enjoyed surfing the Internet (in English). Additionally, she enjoyed watching movies in English in her free time. She thought that the United States and Taiwan had a very good relationship, claiming that nearly everyone that she'd met in the U.S. had been kind to her. She also enjoyed learning English in the ELI, although she felt that she was sometimes too busy to devote as much time to the homework as she'd like to. Grammar was her favorite class.

A181 was a speaker of Arabic. He had chosen to study English because he wanted to enroll as a graduate student in an American university. He said that at first, he had not enjoyed English, because the classes in his home country were "strict" and not very engaging. However, after he had been required to speak English with his boss at work, he came to "love it". He spent time in Michigan where he lived in a dorm with English speakers and had a chance to spend a good deal of time with his roommates; he had also made some American friends that he still kept in touch with via email and instant messaging. While in Pittsburgh he had started meeting with a conversation partner once a week. He also watched American movies in his spare time. He felt pleased with his progress in English and named speaking, listening and reading as his favorite classes. He considered Americans and Saudis to have a good relationship and was generally happy being in the United States.

The two learners with lower motivation and less interaction were C282-3, a Chinese speaker, and A160-3, an Arabic speaker. C282 found learning English to be a struggle, and did not enjoy it; indeed he claimed not to like English very much as a language. He also disliked

most of his teachers and the ELI in general. For instance, when asked about his writing class he responded “I hate my teacher.” Although he did watch some TV in English and read some magazines, he reported that he nearly always spoke Chinese outside of class. On ELI trips, he stayed with Chinese speakers and spoke Chinese with them. When asked about the relationship between Americans and Taiwanese, he replied, “I feel it’s so-so. I’m a foreigner. I don’t think the Americans know where I come from.” (Note that “feeling like a foreigner” is a quality of segregativeness, according to Meisel et al.) The speaker then told a story about being forced to end a badminton game with his Chinese friends so that some Americans could play basketball on a university court. Despite these factors, C282 was motivated to learn in that he needed English to apply to a Master’s program in the United States.

A160-3 claimed to speak mostly Arabic outside of classes. He said he used English only when necessary: on the bus, at the grocery store, etc. He did his homework if he liked the teacher and the activity, but in total he spent less than an hour a week on assignments, and his preferred classes were those in which he didn’t have to study. Learning English was not very fun for him, although it didn’t seem very difficult to him; in fact, he felt that level 5 should be eliminated from the ELI program because it was unnecessary. He planned to enter an American university as an undergraduate as soon as possible and was hoping to score higher on his TOEFL, although he did not plan to study for it.

Recall that Meisel et al. (1981) and Clahsen et al. (1983) predict higher rates of accuracy on variational features for learners with greater motivation and more interaction with native speakers. In order to determine the effect that each learner’s relative amount of interaction and degree of motivation may have had on the acquisition of variational features, each learner’s RSA data was examined for suppliance in obligatory context (i.e., accuracy) of the copula, subjects,

and determiners. (These three variational features were chosen because they are prevalent in the data.) Based on the Multidimensional Model, it is predicted that C127 and A181 (the integrative learners) will show higher rates of suppliance for all three of these elements. C282 and A160 should show lower rates of suppliance for all three elements.

Additionally, an overall rate of error was measured for each learner. These measures were based on Bardovi-Harlig & Bofman (1989), who examined writing samples from advanced-level English learners. Many different kinds of error were noted. For instance, syntactic errors are those of word order, omitted constituents (such as subjects and objects), and errors in sentence combination. Examples of syntactic errors are given below:

(189) The pets they to **they to catch** cats. (C282-3, RSA 1)

(190) [**My father**] Let us growing up and find our future. (A160-3, RSA 3)

It is not clear what structure was intended in (189), so it is counted as a single error. In (190), the subject is omitted, which is also counted as an error.

Morphological errors are those that relate to word form: inflection and derivation. Errors in verb agreement and tense, plural marking, and possessives are included here, as are article errors. Examples of morphological errors are given below.

(191) My father always **advise** me to do the good thing (A160-3, RSA 3)

(192) This city, it has **a** beautiful weather (A181-3, RSA 2)

Finally, lexical-idiomatic errors have to do with word choice, including choice of prepositions.

Examples are given below in (193) and (194).

(193) I want to play badminton to **kill** my health (C127-3, RSA 1)

(194) Younger people is come every weekend to this city to sit **in** the beach.

(A181-3, RSA 3)

These counts were tallied to provide a count of the total number of errors. A single AS-unit may contain numerous errors.

Another prediction of Pienemann (1998) is that segregative learners may fail to develop as well as integrative learners and may fossilize before reaching advanced levels of development. However, none of these learners has fossilized, because three are at an advanced Rapid Profile stage: level 5, and A181 appears to have reached stage 6. Therefore, a simple measure of developmental stage is not enough to distinguish between them. However, it is possible that progress may differ for the two groups in unpredictable ways. Therefore, several types of morphosyntactic complexity are measured for each learner. The prediction is that a complexity measure will show a difference between learners who tend to simplify and those who do not. Note that Pienemann does not predict a relationship between learner orientation and overall level of complexity; I am merely exploring the possibility of a relationship.

Following Bardovi-Harlig & Bofman, as well as standard practice, complexity is first measured as the number of clauses per AS-unit. This measurement is achieved by adding the number of relative clauses, complement clauses, and adjunct clauses, and dividing them by the total number of AS-units for each learner.

One problem with this standard approach is that it relies on the use of adverb and embedded clauses, while ignoring other types of complexity (the use of compound verbs, prepositional phrases, etc.). For instance, by traditional definitions, there is only one clause in the AS-unit below, although the utterance is fairly complex:

- (195) My father always advise me to do the good thing and try to toward [i.e., point] us to the right way. (A160-3, RSA 3)

Thus the utterance in (195) is considered equivalent to the utterance in (196):

(196) I like pets. (C282-3, RSA 1)

To avoid this problem, as an additional measure of complexity, the number of words in the sample will be divided by the number of clauses (AS-units). This measure provides a general idea of the length of each clause. Repetitions, false starts, and fillers will not be counted as words, so as to avoid artificially inflating the word count. Of course, simply using words per clause as a measure of complexity has its drawbacks, as well. For instance, a learner could feasibly list items on a shopping list, which would create a long, very simple clause. A hypothetical example is given in (197):

(197) I bought fruit and vegetables and rice and meat and milk and bread...

For this reason, both measures of complexity are included in this study.

A summary of the measures used to explore the Multidimensional Model are presented in Figure 16.

Measures of Accuracy			Measures of complexity	
Cop. SOC	Det. SOC	Subj. SOC	Clauses per AS- unit	Words per AS-unit

SOC = suppliance in obligatory contexts

Figure 16. Measures used to explore the Multidimensional Model

This completes the explanation of the measures used in this study. In the next section, the research questions will be elaborated.

3.1.6 Research questions

The purpose of this study is to discover whether a developmental measure of proficiency based on morphosyntactic elements is possible. The first step towards this goal is to examine whether the previously proposed developmental paths can account for the production of learners of

varying L1 backgrounds in an intensive English program. If not, it is important to discover whether any path of morphosyntactic development exists that can be used to enhance current proficiency and placement measures.

Five main research questions are posed in this study. They are presented below. In the Discussion section (3.3), these questions will be addressed with respect to the results of the data analysis.

Question 1. Can the path described by Organic Grammar/Minimal Trees account for the morphosyntactic development of these learners in these tasks?

The data will be examined to determine whether learners' morphosyntactic development takes place in the manner outlined by OG. On a related issue, the data will be examined to determine whether learners' morphosyntactic development corresponds to the acquisition of functional categories in Generative grammar (i.e., first IP, then CP). If so, the acquisition of certain features and constructions should be acquired roughly simultaneously. However, there may be non-syntactic reasons (e.g., semantic complexity) that these elements do not pattern together.

In order to make this determination, an implicational table will be created with the results of the analysis described in section 3.1.5.1 above. Coefficients of reproducibility and scalability will be calculated to ascertain if the pattern described by Organic Grammar/Minimal Trees accurately describes these data.

Question 2. Can the path described by Rapid Profile account for the morphosyntactic development of these learners in these tasks?

The data will be examined in a similar manner to that which is used to test Organic Grammar. An implicational table will be created, and a determination will be made as to whether learners' morphosyntactic development takes place in the manner outlined by Rapid Profile.

Question 3. If the paths described by RP and OG do not account for this data, do the morphosyntactic elements appear in a different, but still predictable, order?

The data will be examined to determine whether the additional constructions included in this study are acquired by learners in a particular order, with the goal of discovering whether these constructions can be included in a description of learner development (and therefore in an assessment tool). This determination can be made by adding or removing morphosyntactic elements and sorting the implicational table in ways different from those described by Organic Grammar and Rapid Profile, and then determining whether the coefficients of scalability and reproducibility are greater.

Question 4. To what extent does a learner's first language affect the path of morphosyntactic development?

The data will be examined to determine whether learners with different L1s have different paths of development. This determination can be made first by simple observation; that is, noting whether learners of a particular L1 fail to produce a morphosyntactic element,

overproduce a certain element, or produce it in a different order than other learners. Learners of a certain L1 may also “cluster” at certain points of the implicational table if they share similar characteristics.

If learners of a particular L1 appear to acquire elements in a different order than the others, it is possible to sort learners in the implicational table by L1, and calculate the best fit (i.e., order of elements with the highest coefficient of scalability & reproducibility) for each L1.

Question 5. To what extent does motivation, attitude, and exposure to English affect the path of morphosyntactic development or the accuracy of marking on morphosyntactic elements?

The prediction of both RP and OG is that individual differences should not alter the path of development outlined by the theory, although it may alter the speed with which learners progress through the stages.

The Multidimensional Model predicts superior performance for motivated “integrative” learners on a number of measures (e.g., suppliance of subjects, determiners, and copula). Longitudinal data from four learners, two integrative and two less integrative, will be examined to determine how much their personal experience and behaviors affect their use of these features.

3.1.6 Possible outcomes

There are at least four possible outcomes to the study, which are outlined below:

Outcome One: Organic Grammar and/or Rapid Profile will reliably account for the morphosyntactic development of the ELI learners. Morphological marking develops hand-in-

hand with syntactic development. It may be useful to include additional elements (such as do-support) in the OG assessment profile.

Outcome Two: Organic Grammar and/or Rapid Profile will not reliably account for the morphosyntactic development of the ELI learners; however, adjustments can be made to the ordering of elements to make the system accurate and useable as an assessment tool. Morphosyntactic elements may be added to current proficiency/placement tests.

Outcome Three:

A consistent pattern of syntactic development is found, but learners' use of morphological marking does not develop in a predictable path. Assessment measures based on syntactic development can be used, but morphological marking must be excluded.

Outcome Four:

No consistent pattern of morphosyntactic development can be determined. Therefore, no recommendations can be made for an assessment tool which includes specific morphosyntactic elements as indicators of development.

3.2 RESULTS

3.2.1 Testing Organic Grammar

The results of analysis of the production data were entered into implicational tables. The first table (Table 9) assumes the order of acquisition as presented in Organic Grammar. Because of space constraints, the elements have been abbreviated as shown in Table 8.

Table 8. Abbreviations used in table

Morphosyntactic elements

WO = Word order (SVO)

Sub = no missing subjects

Co = coordination

Pst = 60% past tense marking

Asp = progressive or perfect aspect

RC = relative clauses

CC = complement clauses

Pv = passive

Symbols used to display results

*, ** or -* = emerged feature

/ = no context for this feature

Pro = no missing pronouns

Cop = copula forms besides 'is'

Mod = modals

Agr = 60% agreement on 3rd person singular

B/c = clauses with because

AC = adverb clauses

Ex. = expletive subjects

N = not emerged

Table 9. Implicational Table: Predictions of Organic Grammar

Stud.	WO	Pro	Sub	Cop	Co	Mod	Pst	Agr	Asp	B/c	RC	Ex.	CC	AC	Pv	Pts
A25-2	_*	*	*	*	_*	N	/	/	N	*	N	N	_*	N	N	7
A148-2	*	*	*	_*	*	N	*	N	N	*	N	N	N	N	N	7
K123-2	*	*	*	*	*	N	*	/	N	N	N	N	*	N	N	7
C126-2	*	*	*	_*	_*	N	/	/	N	_*	N	_*	_*	N	N	8
A12-2	*	*	*	N	*	N	*	/	N	N	N	N	_*	_*	N	7
K267-2	*	*	*	_*	*	N	*	/	N	N	N	N	_*	N	N	7
K269-2	*	*	*	*	*	N	*	/	_*	*	N	N	N	_*	N	9
K266-2	*	*	*	**	_*	_*	N	/	N	_*	N	N	N	_*	N	8
K167-3	*	*	*	**	*	_*	N	/	N	*	N	N	*	_*	N	9
K156-2	*	_*	_*	**	_*	_*	*	/	N	N	N	N	_*	_*	N	9
C298-3	*	*	*	*	*	**	/	N	N	N	N	_*	*	*	N	9
C271-2	*	*	*	**	*	_*	*	/	N	_*	N	N	*	*	N	10
S366-2	*	*	*	*	_*	*	/	/	N	_*	*	**	**	N	N	10
A45-2	*	*	*	*	*	_*	*	/	N	*	_*	*	_*	N	N	11
C127-3	*	*	*	**	*	*	*	_*	N	*	N	_*	**	N	N	11
S173-3	*	_*	_*	**	*	*	N	N	_*	*	N	N	*	**	N	10
A338-4	*	*	*	**	*	**	/	N	N	*	**	N	**	*	N	10
K300-3	*	*	*	*	*	_*	*	/	N	*	N	N	_*	*	N	10
A182-4	*	*	*	*	*	*	N	_*	N	N	*	N	**	**	_*	11
K217-4	*	*	*	**	*	**	*	/	N	N	N	*	**	**	N	10
A170-3	*	*	*	**	*	**	*	N	N	_*	_*	_*	N	*	N	11
K111-3	*	*	*	**	*	*	*	/	N	*	N	*	_*	*	N	11
A159-3	*	*	*	*	*	*	*	N	_*	_*	N	N	_*	_*	N	11
K46-3	*	*	*	**	*	N	*	/	_*	_*	N	_*	*	*	N	11
C282-3	*	*	*	**	*	_*	*	/	N	*	*	N	**	_*	_*	12
S391-4	*	_*	_*	**	*	**	/	/	*	*	**	N	**	*	N	11
K101-3	*	*	*	**	*	**	*	/	N	*	N	*	*	*	N	11
A279-4	*	*	*	**	*	**	*	/	N	*	*	N	**	*	N	11
A157-3	*	*	*	**	*	*	*	/	N	*	**	*	*	_*	N	12
S130-3	*	_*	_*	**	*	**	/	_*	*	*	_*	N	**	_*	N	12
S362-2	*	*	*	*	*	**	*	/	N	*	_*	*	**	_*	N	12
T397-5	*	*	*	**	*	*	*	/	_*	*	*	*	**	N	_*	13
C177-3	*	*	*	**	*	*	*	*	N	*	N	_*	**	*	N	12
C84-4	*	*	*	*	*	**	*	/	N	*	**	_*	**	**	N	12
C278-4	*	*	*	*	*	**	N	N	_*	*	*	_*	**	**	_*	13
C201-4	*	*	*	**	*	*	*	/	*	*	_*	N	**	*	N	12
C301-3	*	*	*	**	*	**	N	_*	N	*	*	*	**	**	N	12
K320-4	*	*	*	**	*	**	N	/	*	*	*	_*	**	**	N	12
C118-4	*	*	*	**	N	**	*	*	N	*	N	**	**	*	_*	12
S383-4	*	_*	_*	**	*	**	/	*	**	*	*	*	**	_*	N	13
A160-3	*	*	*	*	*	*	*	*	_*	*	_*	N	_*	*	N	13
A163-3	*	*	*	**	*	_*	*	N	**	*	_*	_*	**	*	N	13
K275-4	*	*	*	*	*	*	*	/	_*	_*	_*	**	**	**	N	13
A199-4	*	*	*	**	*	_*	*	/	*	*	_*	*	**	**	N	13
A161-3	*	*	*	**	*	*	*	_*	*	*	_*	_*	**	**	N	14
K276-4	*	*	*	**	*	**	*	_*	_*	*	N	N	**	**	_*	13
S100-5	*	*	*	**	*	**	*	/	_*	*	**	*	**	**	**	14
J274-5	*	*	*	**	*	**	*	/	N	*	**	_*	**	**	_*	13

Notice that the line separating the “emerged” from the “non-emerged” side is drawn high in the table (with the exception of passive), so that the majority of the table is “emerged.” The line was drawn high in the table because it created the minimum amount of error; despite the odd appearance of the table, it produces the highest possible Coefficient of Reproducibility.

The fact that the line must be drawn high (again, except for passive) points to another interesting fact about the table. Notice that the students who tested into level 2 by traditional means (the Michigan Test and a writing sample) cluster at the top of the table. That is, they predictably fail to produce many of the morphosyntactic forms analyzed here. However, the students who tested into levels 3, 4 and 5 are scattered throughout the rest of the table; they do not cluster together. This observation will be explored further in the Discussion.

There are 94 errors in the table (N’s on the left and *’s on the right of the line), out of 682 total answers (excluding 38 non-applicable responses, i.e., where there was no context for inflection). Therefore, the C. of R. is calculated as follows:

$$(198) \text{ C. of R.} = 1 - (94/682) = .86$$

This result is lower than is required to indicate a predictable table. In other words, the results of individual students cannot be reliably predicted based on this order of acquisition.

Further calculations were performed: the Minimal Marginal Reproducibility (MMR) is the number of emerged responses divided by the total, as in (199), while the Percentage Improvement of Reproducibility (PIR) is the difference between the MMR and the C of R. It is calculated in (200).

$$(199) \text{ MMR} = 521/682 = .76$$

$$(200) \text{ PIR} = .86 - .76 = 0.1$$

Finally, the Coefficient of Scalability is calculated based on the above figures. $C\ of\ S = PIR / (1 - MMR)$

$$(201) \ C\ of\ S = .1 / (1 - .76) = .42$$

This number is obviously lower than the .6 to .65 that indicates a scalable table (i.e., a real developmental pattern).

It is also possible to test the reproducibility of individual test items. This information indicates whether students' responses on a particular element can be predicted. The C of R has been calculated for each of the items in the chart above; those data are presented below in Table 10. Results for Word Order, Subjects and Pronouns are not included because they are all "1"; i.e., there are no errors at all because those elements had emerged for every student.

Table 10. Coefficients of reproducibility for elements in Organic Grammar

Element	C of R	Element	C of R
Copula	.98*	Coordination	.98*
Complement Cl.	.79	Past tense	.75
Because Cl.	.81	Agreement	.67
Modals	.98*	Adverb Cl.	.83
Expletives	.73	Relative Cl.	.77
Aspect	.58	Passive	.92*

The use of the copula and coordination are reproducible in this table, but this fact is only trivially true, since only one learner does not yet have emerged coordination and only one lacks copula forms beyond *is*; thus it is a very simple matter to predict that all learners will have copula and coordination. Given that a number of learners fail to produce modals and passive, it is more interesting to see that these two elements also have a high coefficient. That means that learners'

use of modals and passive is highly predictable by this table. However, none of the other elements emerge in a predictable fashion based on this table. Note that the low coefficients do not indicate that the elements never emerge in a predictable way; they only indicate that they do not emerge predictably based on this order of acquisition.

In Table 11, the elements of Organic Grammar have been reordered to create a table with maximum predictability and scalability. An additional element, do-support (Do/s), has also been added; it is predicted by Minimal Trees to emerge with IP features, but it is left out of the original Organic Grammar table.

Table 11. OG implicational table revised for greatest scalability and reproducibility

Student	WO	Pro	Sub	Cop	Co.	CC	Pst	B/c	Agr	Mod	AC	Do/s	Ex.	RC	Asp	Pv	Pts
K123-2	*	*	*	*	*	*	*	N	/	N	N	N	N	N	N	N	7
A12-2	*	*	*	N	*	*-	*	N	/	N	*-	N	N	N	N	N	7
K267-2	*	*	*	*-	*	*-	*	N	/	N	N	**	N	N	N	N	8
A148-2	*	*	*	*-	*	N	*	*	N	N	N	N	N	N	N	N	7
A25-2	*-	*	*	*	*-	*-	/	*	/	N	N	N	N	N	N	N	7
C126-2	*	*	*	*-	*-	*-	/	*-	/	N	N	N	*-	N	N	N	8
K269-2	*	*	*	*	*	N	*	*	/	N	*-	N	N	N	*-	N	9
S366-2	*	*	*	*	*-	**	/	*-	/	*	N	N	**	*	N	N	10
K266-2	*	*	*	**	*-	N	N	*-	/	*-	*-	*	N	N	N	N	9
K167-2	*	*	*	**	*	*	N	*	/	*-	*-	N	N	N	N	N	9
K156-2	*	*-	*-	**	*-	*-	*	N	/	*-	*-	N	N	N	N	N	9
C271-2	*	*	*	**	*	*	*	*-	/	*-	*-	N	N	N	N	N	10
C298-3	*	*	*	*	*	/	N	N	N	**	*	**	*-	N	N	N	10
A45-2	*	*	*	*	*	*-	*	*	/	*-	N	N	*	*-	N	N	11
C127-3	*	*	*	**	*	**	*	*	*-	*	N	N	*-	N	N	N	11
C118-4	*	*	*	**	N	**	*	*	*	**	*	N	**	N	N	*-	11
A170-2	*	*	*	**	*	N	*	*-	N	**	*	N	*-	*-	N	N	11
K111-3	*	*	*	**	*	*-	*	*	/	*	*	N	*	N	N	N	11
A157-3	*	*	*	**	*	*	*	*	/	*	*-	N	*	**	N	N	11
S173-3	*	*-	*-	**	*	*	N	*	N	*	**	*-	N	N	*-	N	11
A338-4	*	*	*	**	*	**	/	*	N	**	*	**	N	**	N	N	11
K300-3	*	*	*	*	*	*-	*	*	/	*-	*	*-	N	N	N	N	11
A182-4	*	*	*	*	*	**	N	N	*-	*	**	*-	N	*	N	*-	11
K217-4	*	*	*	**	*	**	*	N	/	**	**	*	*	N	N	N	11
A159-3	*	*	*	*	*	*-	*	*-	N	*	*-	*-	N	N	*-	N	12
S130-3	*	*-	*-	**	*	**	/	*	*-	**	*-	N	N	*-	*	N	12
K46-3	*	*	*	**	*	*	*	*-	/	N	*	*	*-	N	*-	N	12
C282-3	*	*	*	**	*	**	*	*	/	*-	*-	*	N	*	N	*-	12
S391-4	*	*-	*-	**	*	**	/	*	/	**	*	*	N	**	*	N	12
K101-3	*	*	*	**	*	*	*	*	/	**	*	*-	*	N	N	N	12
A279-4	*	*	*	**	*	**	*	*	/	**	*	*-	N	*	N	N	12
K276-4	*	*	*	**	*	**	*	*	*-	**	**	*	N	N	*-	*-	13
C177-3	*	*	*	**	*	**	*	*	*	*	*	*-	*-	N	N	N	13
C84-4	*	*	*	*	*	**	*	*	/	**	**	*-	*-	**	N	N	13
C301-3	*	*	*	**	*	**	N	*	*-	**	**	**	*	*	N	N	13
S362-2	*	*	*	*	*	**	*	*	/	**	*-	N	*	*-	N	N	12
J274-5	*	*	*	**	*	**	*	*	/	**	**	*-	*-	**	N	*-	13
S383-4	*	*-	*-	**	*	**	/	*	*	**	*-	N	*	*	**	N	13
K320-4	*	*	*	**	*	**	N	*	/	**	**	**	*-	*	*	N	13
C278-4	*	*	*	*	*	**	N	*	N	**	**	*	*-	*	*-	*-	13
T397-5	*	*	*	**	*	**	*	*	/	*	N	N	*	*	*-	*-	12
A160-3	*	*	*	*	*	*-	*	*	*	*	*	*	N	*-	*-	N	14
A163-3	*	*	*	**	*	**	*	*	N	*-	*	*	*-	*-	**	N	14
K275-4	*	*	*	*	*	**	*	*-	/	*	**	*	**	*-	*-	N	14
A199-4	*	*	*	**	*	**	*	*	/	*-	**	*-	*	*-	*	N	14
C201-4	*	*	*	**	*	**	*	*	/	*	*	*-	N	*-	*	N	13
A161-3	*	*	*	**	*	**	*	*	*-	*	**	*-	*-	*-	*	N	15
S100-5	*	*	*	**	*	**	*	*	/	**	**	*-	*	**	*-	**	15

Notice that the line separating the “emerged” from the “non-emerged” side does not need to be as high in the table as in Table 10 to produce the highest Coefficient of Reproducibility (i.e., minimum errors). However, it remains fairly far to the right, indicating that most of the students have a majority of these features already in their morphosyntax. Notice that the same effect occurs in this table as in the earlier one: students who tested into level 2 cluster towards the top of the chart, while those who tested into levels 3, 4 and 5 are scattered throughout.

There are 75 errors in the table (N’s on the left and *’s on the right of the line), out of 730 total answers (excluding 38 non-applicable responses). Therefore, the C. of R. is calculated as follows:

$$(202) \text{ C. of R.} = 1 - (75/730) = .90$$

This result is .9, the minimum number that is required to indicate a predictable table. However, there are three measures which fail to differentiate between learners, i.e., everyone has the same result: word order, pronouns, and subjects. If these measures are removed, the Coefficient of Reproducibility drops to .87. This number is below .9 but approaches significance.

The Coefficient of Scalability is now computed. First, the Minimal Marginal Reproducibility (MMR) is the number of emerged responses divided by the total, as in (203), while the Percentage Improvement of Reproducibility (PIR) is the difference between the MMR and the C of R. It is calculated in (204).

$$(203) \text{ MMR} = 549/730 = .75$$

$$(204) \text{ PIR} = .9 - .75 = .15$$

Finally, the Coefficient of Scalability is calculated based on the above figures.

$$(205) \text{ C of S} = .15 / (1 - .75) = .6$$

This number indicates a scalable table (i.e., a real developmental pattern), and it is obviously an improvement over the previous measure. If the three measures are removed which fail to distinguish between learners—word order, pronouns, and subject—the Coefficient of Scalability drops to .58.

The coefficients of reproducibility have been calculated for the elements in this table. Note that many of the figures are different from those in Table 10, because the number of errors has been changed for each element.

Table 12. Coefficients of reproducibility for elements in revised OG table

Element	C of R	Element	C of R
Copula	.98*	Coordination	.98*
Complement Cl.	.94*	Past tense	.83
Because Cl.	.88(*)	Agreement	.61
Modals	.98*	Adverb Cl.	.90*
Expletives	.71	Relative Cl.	.79
Aspect	.56	Do-support	.88(*)
Passive	.85		

As in Table 10, both the copula and coordination have very high coefficients, but this is to be expected because only one learner fails to produce each of those structures. Modals again have a high C of R, indicating that they emerge predictably for learners. Four other elements—*because* clauses, do-support, and complement and adverb clauses—also have coefficients that reach or nearly reach .9. This figure indicates that it is possible to predict students’ performance on these elements, albeit in a different order than is predicted by OG.

A final implicational table was created to examine the outcome if a higher SOC was used to measure past tense suppliance and agreement marking. Note that because few learners used

third person singular forms, little can be said about agreement. However, most learners used enough tokens of past tense marking to provide meaningful results for tense. In Table 13, the “cutoff point” is raised to 80% suppliance for inflectional marking.

Table 13. Implicational Table revised with cutoff points of 80% for inflection

Stud.	WO	Pro	Sub	Cop	Co	CC	B/c	Mod	AC	Pst 80	Do/s	Ex	RC	Asp	Agr 80
K123-2	*	*	*	*	*	*	N	N	N	N	N	N	N	N	/
A148-2	*	*	*	*_	*	N	*	N	N	N	N	N	N	N	/
A25-2	*_	*	*	*	*_	*_	*	N	N	/	N	N	N	N	/
A12-2	*	*	*	N	*	*_	N	N	*_	N	N	N	N	N	/
C126-2	*	*	*	*_	*_	*_	*_	N	N	/	N	*_	N	N	/
K267-2	*	*	*	*_	*	*_	N	N	N	N	**	N	N	N	/
K26902	*	*	*	*	*	N	*	N	*_	*	N	N	N	*_	/
K266-2	*	*	*	**	*_	N	*_	*_	*_	N	*	N	N	N	/
K167-2	*	*	*	**	*	*	*	*_	*_	N	N	N	N	N	/
K156-2	*	*_	*_	**	*_	*_	N	*_	*_	*	N	N	N	N	/
C271-2	*	*	*	**	*	*	*_	*_	*	N	N	N	N	N	/
S366-2	*	*	*	*	*_	**	*_	*	N	/	N	**	*	N	/
C298-3	*	*	*	*	*	*	N	**	*	/	**	*_	N	N	N
A45-2	*	*	*	*	*	*_	*	*_	N	N	N	*	*_	N	/
C127-3	*	*	*	**	*	**	*	*	N	*	N	*_	N	N	/
C118-4	*	*	*	**	N	**	*	**	*	*	N	**	N	N	*
A170-A	*	*	*	**	*	N	*_	**	*	*	N	*_	*_	N	/
K111-3	*	*	*	**	*	*_	*	*	*	*	N	*	N	N	/
A157-3	*	*	*	**	*	*	*	*	*_	*	N	*	**	N	/
S173-3	*	*_	*_	**	*	*	*	*	**	N	*_	N	N	*_	N
A338-4	*	*	*	**	*	**	*	**	*	/	**	N	**	N	N
K300-3	*	*	*	*	*	*_	*	*_	*	N	*_	N	N	N	/
A182-4	*	*	*	*	*	**	N	*	**	N	*_	N	*	N	/
K217-4	*	*	*	**	*	**	N	**	**	*	*	*	N	N	/
A159-3	*	*	*	*	*	*_	*_	*	*_	N	*_	N	N	*_	N
S130-3	*	*_	*_	**	*	**	*	**	*_	/	N	N	*_	*	/
K46-3	*	*	*	**	*	*	*_	N	*	*	*	*_	N	*_	/
S362-2	*	*	*	*	*	**	*	**	*_	*	N	*	*_	N	/
C282-3	*	*	*	**	*	**	*	*_	*_	*	*	N	*	N	/
S391-4	*	*_	*_	**	*	**	*	**	*	/	*	N	**	*	/
K101-3	*	*	*	**	*	*	*	**	*	*	*_	*	N	N	/
A279-4	*	*	*	**	*	**	*	**	*	N	*_	N	*	N	/
T397-5	*	*	*	**	*	**	*	*	N	*	N	*	*	*_	/
K276-4	*	*	*	**	*	**	*	**	**	N	*	N	N	*_	/
C177-3	*	*	*	**	*	**	*	*	**	N	*_	*_	N	N	N
C84-4	*	*	*	*	*	**	*	**	**	*	*_	*_	**	N	/
C301-3	*	*	*	**	*	**	*	**	**	N	**	*	*	N	/
J274-5	*	*	*	**	*	**	*	**	**	*	*_	*_	**	N	/
S383-4	*	*_	*_	**	*	**	*	**	*_	/	N	*	*	**	*
C278-4	*	*	*	*	*	**	*	**	**	N	*	*_	*	*_	N
C201-4	*	*	*	**	*	**	*	*	*	N	*_	N	*_	*	/
K320-4	*	*	*	**	*	**	*	**	**	N	**	*_	*	*	/
A160-3	*	*	*	*	*	*_	*	*	*	*	*	N	*_	*_	N
A163-3	*	*	*	**	*	**	*	*_	*	N	*	*_	*_	**	N
K275-4	*	*	*	*	*	**	*_	*	**	*	*	**	*_	*_	/
A199-4	*	*	*	**	*	**	*	*_	**	N	*_	*	*_	*	/
S100-5	*	*	*	**	*	**	*	**	**	N	*_	*	**	*_	/
A161-3	*	*	*	**	*	**	*	*	**	*	*_	*_	*_	*	/

Both the tense and agreement column are moved further to the right of the table to minimize errors. However, the number of errors in this table is higher than in the table with the 60% cutoff point, because of the large number of N's even at advanced levels. Therefore, raising the cutoff point of SOC for inflection makes the table less reproducible and less scalable.

Finally, in order to assess whether eliminating morphological elements improves the reproducibility and scalability of the table, measures of tense and agreement marking were removed from the table. The result is the implicational table shown in Table 14.

Table 14. Revised OG table with morphological measures removed

Stud.	WO	Pro	Sub	Cop	Co	Cmp	B/c	Mod	AC	Do/s	Ex	RC	Asp	Pv
K123-2	*	*	*	*	*	*	N	N	N	N	N	N	N	N
A148-2	*	*	*	*_	*	N	*	N	N	N	N	N	N	N
A25-2	*_	*	*	*	*_	*_	*	N	N	N	N	N	N	N
A12-2	*	*	*	N	*	*_	N	N	*_	N	N	N	N	N
C126-2	*	*	*	*_	*_	*_	*_	N	N	N	*_	N	N	N
K267-2	*	*	*	*_	*	*_	N	N	N	**	N	N	N	N
K269-2	*	*	*	*	*	N	*	N	*_	N	N	N	*_	N
K266-2	*	*	*	**	*_	N	*_	*_	*_	*	N	N	N	N
K167-2	*	*	*	**	*	*	*	*_	*_	N	N	N	N	N
K156-2	*	*_	*_	**	*_	*_	N	*_	*_	N	N	N	N	N
C271-2	*	*	*	**	*	*	*_	*_	*	N	N	N	N	N
S366-2	*	*	*	*	*_	**	*_	*	N	N	**	*	N	N
C298-3	*	*	*	*	*	*	N	**	*	**	*_	N	N	N
A45-2	*	*	*	*	*	*_	*	*_	N	N	*	*_	N	N
C127-3	*	*	*	**	*	**	*	*	N	N	*_	*_	N	N
C118-4	*	*	*	**	N	**	*	**	*	N	**	N	N	*_
A170-A	*	*	*	**	*	N	*_	**	*	N	*_	*_	N	N
K111-3	*	*	*	**	*	*_	*	*	*	N	*	N	N	N
A157-3	*	*	*	**	*	*	*	*	*_	N	*	**	N	N
S173-3	*	*_	*_	**	*	*	*	*	**	*_	N	N	*_	N
A338-4	*	*	*	**	*	**	*	**	*	**	N	**	N	N
K300-3	*	*	*	*	*	*_	*	*_	*	*_	N	N	N	N
A182-4	*	*	*	*	*	**	N	*	**	*_	N	*	N	*_
A159-3	*	*	*	*	*	*_	*_	*	*_	*_	N	N	*_	N
S130-3	*	*_	*_	**	*	**	*	**	*_	N	N	*_	*	N
C282-3	*	*	*	**	*	**	*	*_	*_	*	N	*	N	N
K276-4	*	*	*	**	*	**	*	**	**	*	N	N	*_	N
K217-4	*	*	*	**	*	**	N	**	**	*	*	N	N	N
C177-3	*	*	*	**	*	**	*	*	*	*_	*_	N	N	N
K101-3	*	*	*	**	*	*	*	**	*	*_	*	N	N	N
K46-3	*	*	*	**	*	*	*_	N	*	*	*_	N	*_	N
T397-5	*	*	*	**	*	**	*	*	N	N	*	*	*_	N
C84-4	*	*	*	*	*	**	*	**	**	*_	*_	**	N	N
C301-3	*	*	*	**	*	**	*	**	**	**	*	*	N	*_
J274-5	*	*	*	**	*	**	*	**	**	*_	*_	**	N	N
A279-4	*	*	*	**	*	**	*	**	*	*_	N	*	N	*_
S362-2	*	*	*	*	*	**	*	**	*_	N	*	*_	N	*_
S391-4	*	*_	*_	**	*	**	*	**	*	*	N	**	*	N
K320-4	*	*	*	**	*	**	*	**	**	**	*_	*	*	N
A160-3	*	*	*	*	*	*_	*	*	*	*	N	*_	*_	N
A163-3	*	*	*	**	*	**	*	*_	*	*	*_	*_	**	N
K275-4	*	*	*	*	*	**	*_	*	**	*	**	*_	*_	N
A199-4	*	*	*	**	*	**	*	*_	**	*_	*	*_	*	N
S100-5	*	*	*	**	*	**	*	**	**	*_	*	**	*_	N
S383-4	*	*_	*_	**	*	**	*	**	*_	N	*	*	**	N
C201-4	*	*	*	**	*	**	*	*	*	*_	N	*_	*	*_
C278-4	*	*	*	*	*	**	*	**	**	*	*_	*	*_	*_
A161-3	*	*	*	**	*	**	*	*	**	*_	*_	*_	*	**

There are 60 errors in the table (N's on the left and *'s on the right of the line), out of 624 total answers. Therefore, the C. of R. is calculated as follows:

$$(205) \text{ C. of R.} = 1 - (60/672) = .91$$

This result is .91, a very slight improvement over the results for the table with morphological elements included.

The Coefficient of Scalability is now computed. First, the Minimal Marginal Reproducibility (MMR) is the number of emerged responses divided by the total, as in (206), while the Percentage Improvement of Reproducibility (PIR) is the difference between the MMR and the C of R. It is calculated in (207).

$$(206) \text{ MMR} = 506/672 = .75$$

$$(207) \text{ PIR} = .91 - .75 = .16$$

Finally, the Coefficient of Scalability is calculated based on the above figures.

$$(208) \text{ C of S} = .16 / (1 - .75) = .64$$

Again, this number represents a small improvement over previous measures.

These results and manipulations will be discussed further in section 4. For now, it should be noted that the implicational table with the best fit does not follow the original predictions of Organic Grammar.

3.2.2 Testing Rapid Profile

An implicational table was created for the Rapid Profile measures. The morphosyntactic elements in Rapid Profile are somewhat different than those in Organic Grammar, so the abbreviations are presented in Table 15. The implicational table itself is given in Table 16.

Table 15. Abbreviations and symbols used in Rapid Profile table

Morphosyntactic elements	
WO = word order	Plur = regular plural (s)
Ing = verbs with –ing	ed = regular past tense-marked verbs
Poss pro = possessive pronouns	Adv. 1 st = adverb (or prepositional phrase, etc.) as first element in a sentence
Obj pro = object pronouns	3 rd –s = 3 rd person singular inflection (s)
Cncl Aux = wh-complement clauses	
Symbols used to display results	
*, ** or -* = emerged feature	N = not emerged
/ = no context for this feature	

Table 16. Implicational table: Predictions of Rapid Profile

Student	WO	Plur	-ing	-ed	Poss pro	Adv 1 st	Obj pro	3 rd -s	Cncl Aux	Pts
A148-2	*	N	*_	/	*	N	*	N	N	4
A25-2	*	*	N	*_	*	N	N	/	N	4
K123-2	*	*	N	*	*	N	N	/	N	4
K266-2	*	*	*	*	*	N	N	/	N	5
C126-2	*	*	N	/	N	*	N	/	N	3
A170-2	*	*	N	*	*	*	N	N	N	5
S366-2	*	*	N	*_	*	*	N	/	N	5
S362-2	*	*	N	*	*	*	N	/	N	5
A45-2	*	*	*_	*_	N	*	N	/	N	5
A159-3	*	*	*	*	*	*	N	N	N	6
A338-4	*	*	*	/	*	*	N	*	N	6
A199-4	*	*	*	*	*	*	N	/	N	6
K111-3	*	*	*_	*	*	*	N	/	N	6
T397-5	*	*	*	*	*	*	N	/	N	6
K101-3	*	*	*	*	*	*	N	*_	N	7
J84-4	*	*	*	*	*	*	N	*_	N	7
K167-3	*	N	N	*	*	*	*_	/	N	5
K156-2	*	*	N	*_	*	*	*	/	N	6
A12-2	*	*	*	*_	N	*	*_	/	N	6
K267-2	*	*	*	N	*	*	*_	/	N	6
K300-3	*	*	N	*	*	*	*	/	N	6
C282-3	*	*	N	*	*	*	*	N	N	6
C298-3	*	*	*_	N	*	*	*	N	N	6
S100-5	*	*	N	*	*	*	*	/	N	6
C271-2	*	*	N	N	*	*	*	*_	N	6
A163-3	*	*	*	*	*	*	*	N	N	7
A279-4	*	*	*	*	*	*	*	/	N	7
K269-2	*	*	*	*	*	*	*	/	N	7
K46-3	*	*	*_	*	*	*	*_	/	N	7
K217-4	*	*	N	*	*	*	*	/	*	7
K276-4	*	*	*	*	*	*	*	N	N	7
C127-3	*	*	*_	*	*	*	*_	N	N	7
C201-4	*	*	*	*	*	*	*	/	N	7
S391-4	*	*	*	/	*	*	*_	N	*	7
S130-3	*	*	*	/	*	*	*	*	N	7
A160-3	*	*	*	*_	*	*	*	*	N	8
A161-3	*	*	*	*	*	*	*	*	N	8
A182-4	*	*	*	*	*	*	*	*_	N	8
K320-4	*	*	*_	*	*	*	*	*_	N	8
C177-3	*	*	*	*	*	*	*	*	N	8
C301-3	*	*	*_	*	*	*	*	*_	N	8
C118-4	*	*	*	*	*	*	*_	*	N	8
S173-3	*	*	*	*_	*	*	*	*	N	8
C278-4	*	*	*	N	*	*	*	*	*	8
A157-3	*	*	*	*	*	*	N	*	*	8
S383-4	*	*	*	/	*	*	*_	*	*	8
J274-5	*	*	*	*	*	*	*	/	*	8
K275-4	*	*	*	*	*	*	*	*_	*	9

Notice that the line separating the “emerged” from the “non-emerged” side extends almost all the way to the bottom of the table; however, it is on the far right, similar to the separating lines in the Organic Grammar tables. That is, most of the learners have most of the elements in the table in their linguistic system. The result is that, as with the Organic Grammar tables, most of the elements only distinguish between beginning levels. When students are intermediate to advanced, nearly all of the elements on the table have emerged. Another similarity to the OG table is that students who placed into level 2 appear uniformly at the top of the table, while levels 3, 4 and 5 are scattered throughout the rest.

Another interesting finding is that the order of learners in terms of development is different for Organic Grammar and Rapid Profile. That is, the students that score the highest on the Organic Grammar measures are not the same students who score the highest on the Rapid Profile. For this reason, it is not possible to combine the two measures into a single one.

There are 30 errors in the table (N’s on the left and *’s on the right), out of 404 total answers (excluding 28 non-applicable responses). Therefore, the C of R is calculated as follows:

$$(209) \text{ C. of R.} = 1 - (30/404) = .93$$

This result is .93, indicating that the results for individuals can be predicted based on the table. Furthermore, even if the “Word Order” results are removed (which provide no information about individuals in these data, since all learners have SVO word order), the C of R calculates as .91, still a significant result.

In order to compute the Coefficient of Scalability, we must first calculate the Minimal Marginal Reproducibility (MMR), which is the number of emerged responses divided by the total, as in (210). The Percentage Improvement of Reproducibility (PIR) is the difference between the MMR and the C of R. It is calculated in (211).

$$(210) \text{ MMR} = 312/404 = .77$$

$$(211) \text{ PIR} = .93 - .77 = .16$$

Finally, the Coefficient of Scalability is calculated based on the above figures. C of S = PIR / (1 – MMR)

$$(212) \text{ C of S} = .16 / (1 - .77) = .7$$

This number is greater than .6 to .65 and therefore indicates a scalable table (i.e., a real developmental pattern). Furthermore, if the results for word order are removed, the C of S calculates as 6.5, which is still a scalable table. Reordering the table does not produce a higher C of R or C of S.

The individual coefficients of reproducibility have been calculated for the elements in the table; the results are given below.

Table 17. Coefficients of reproducibility for elements in the Rapid Profile table

Element	C of R	Element	C of R
Plural	.96*	-ing	.73
Past	.90*	Possess. pron.	.93*
Adverb 1 st	1.0*	Object pron.	.95*
3 rd sing –s	.89(*)	Cancel Aux 2 nd	.95*

Nearly all of the elements reach .9 or nearly reach .9 when the chart based on Rapid Profile is used. The only element that fails to meet the criterion is *-ing*, which does not appear in a predictable way in spontaneous production data.

3.2.3 Other measures

Following Young-Scholten, Vainikka & Ijuin, I counted the number of AS-units in each learner's data sample and measured vocabulary as the number of unique nouns and total nouns per sample. That information is provided in Table 18. This data is provided as a reference; it was not included in the analysis.

Table 18. Sample size and vocabulary measures for learners

Student	Size in AS units	Unique nouns/ total nouns
A45-2	17	21/24
A12-2	19	17/20
A148-2	25	11/25
A25-2	13	9/13
K123-2	25	16/29
K156-2	27	20/32
K269-2	15	15/22
K267-2	22	13/34
K266-2	22	19/37
C271-2	25	18/29
C126-2	13	7/11
S366-2	16	22/38
S362-2	18	14/30
A157-3	29	38/57
A159-3	26	24/35
A160-3	33	22/36
A161-3	53	23/53
A163-3	45	35/65
A170-3	34	30/46
K46-3	26	25/40
K101-3	38	27/54
K111-3	50	39/76
K167-3	31	23/30
K300-3	18	23/30
C282-3	36	41/66
C177-3	39	30/36
C298-3	35	30/58
C301-3	34	34/69
C127-3	39	39/61
S173-3	35	30/49
S130-3	51	57/112
A338-4	28	36/64
A279-4	19	42/64
A199-4	38	37/54
A182-4	27	46/66
K217-4	31	26/43
L320-4	30	27/41
L275-4	54	28/58
K276-4	44	29/66
C84-4	47	44/91
C118-4	39	40/63
C278-4	49	33/72
C201-4	31	34/60
S391-4	27	41/77
S383-4	29	44/73
S100-5	29	47/90
J274-5	31	60/98
T397-5	38	47/76

The next section presents the results from the analysis of interviews.

3.2.4 Interview results

3.2.4.1 Overall development

The learners' data was examined for morphosyntactic features. That information is presented below, in Table 19. Learners that are listed as (I) are considered to be integrative based on the interviews, while learners listed with (LI) are less integrative.

Table 19. Morphosyntactic features of integrative and less integrative learners

Stud.	Lev.	WO	Co	Sub	Adv1	Mod	CC	Cop	Pst	B/c	Agr.	AC	RC	Ex	Asp	Pv
C127 (I)	3	*	*	*	*	*	**	*	*	**	/	N	N	_*-	N	N
C127	4	*	*	*	**	**	**	*	_*-	*	/	*	N	N	N	N
C127	5	*	*	*	**	*		*	_*-	*	/	*	*	N	N	N
C282 (LI)	3	*	*	*	*	*	**	*	*	*	/	_*-	*	N	N	_*-
C282	4	*	*	*	**	**	**	*	/	_*-	/	**	N	_*-	N	N
C282	5	*	*	*	*	*	_*-	N	/	_*-	/	*	N	N	N	N
A181 (I)	3	*	*	_*-	*	*	**	*	_*-	**	_*-	*	N	N	_*-	N
A181	4	*	*	_*-	**	*	**	_*-	_*-	**	/	**	N	*	N	N
A181	5	*	*	*	**	**	**	N	**	*	/	*	N	N	N	N
A160 (LI)	3	*	*	*	*	*	N	_*-	**	*	_*-	*	_*-	N	*	N
A160	4	*	*	*	N	*	**	*	/	_*-	N	N	*	N	N	N
A160	5	*	*	*	*	**	_*-	N	/	*	/	*	**	N	*	N

Development appears to progress along similar lines for all four of the learners. That is, they all appear to have appropriate word order, coordination, subjects, sentences with non-initial subjects, modals, complement clauses, copula forms other than 'is', past tense marking, because-

clauses, and adverb clauses. Not yet emerged (or only partially emerged) are passives, aspect, expletive subjects, and relative clauses (except for A160, who produces RCs).

3.2.4.2 Error measures

Learners' data was examined for three variational features: copula, subject pronouns, and determiners. Based on the Multidimensional Model, it is predicted that integrative learners (marked with (I) in the tables) will have a higher rate of suppliance than less integrative learners (marked with (LI) in the tables).

First, suppliance in obligatory context of the copula is presented in Table 20.

Table 20. Copula suppliance of integrative and less integrative learners

Learner	Level 3	Level 4	Level 5
C127 (I)	100% (7/7)	54% (7/13)	80% (8/10)
C282 (LI)	73% (8/11)	100% (5/5)	100% (7/7)
A181 (I)	100% (4/4)	100% (5/5)	100% (5/5)
A160 (LI)	82% (9/11)	100% (4/4)	80% (4/5)

A graph presenting the same information is given in Figure 17.

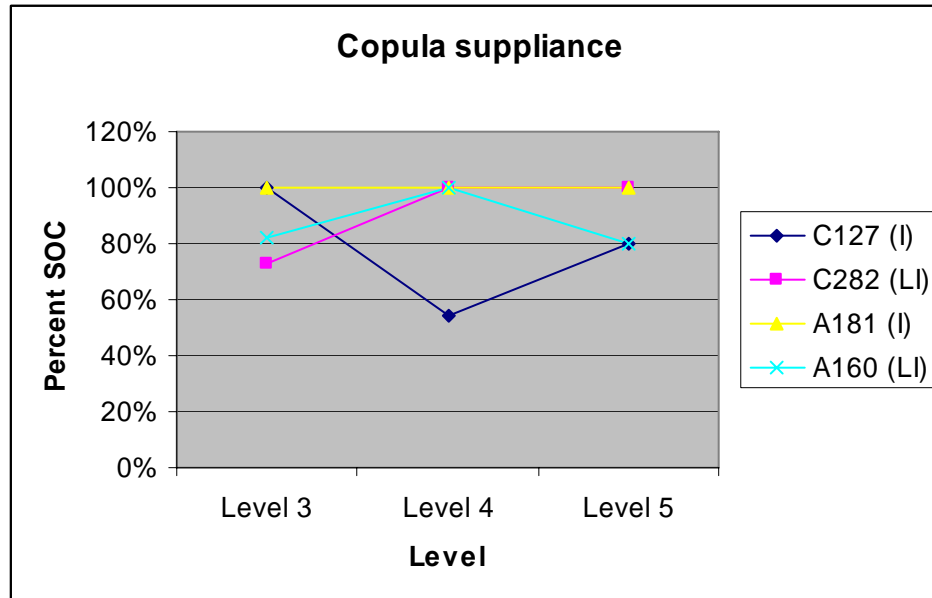


Figure 17. Copula suppliance of integrative and less integrative learners

Note that because the number of contexts is relatively small, a small difference in suppliance creates a dramatic effect on the graph. Generally the learners exhibit relatively high suppliance, but C127 appears to omit more copulas than the others. This is not predicted by the Multidimensional Model, since C127 is integrative. The other learner categorized as integrative, A181, omits no copulas at all, in any of the three levels.

Next, the rates of suppliance for subject pronouns is presented in Table 21.

Table 21. Subject suppliance for integrative and less integrative learners

Learner	Level 3	Level 4	Level 5
C127 (I)	100% (43/43)	100% (52/52)	100% (39/39)
C282 (LI)	91% (38/42)	97% (37/38)	100% (14/14)
A181 (I)	100% (41/41)	100% (26/26)	100% (18/18)
A160 (LI)	100% (34/34)	100% (19/19)	100% (25/25)

Note that subject omission is quite rare in these learners. The only exception is C282, who omits 9% of subjects in level 3. By the time he reaches level 5, however, he supplies 100% of subjects. This measure does not serve to distinguish the integrative learners from the less integrative learners.

Rates of determiner suppliance are presented next. Table 22 shows suppliance in obligatory contexts for learners who are integrative and less integrative.

Table 22. Determiner suppliance in integrative and less integrative learners

Learner	Level 3	Level 4	Level 5
C127 (I)	72% (16/23)	89% (11/15)	90% (16/20)
C282 (LI)	81% (22/27)	88% (12/18)	84% (25/30)
A181 (I)	81% (18/25)	67% (39/44)	83% (27/30)
A160 (LI)	70% (21/26)	73% (22/25)	80% (8/10)

The same information is provided in graph form below, in Figure 18.

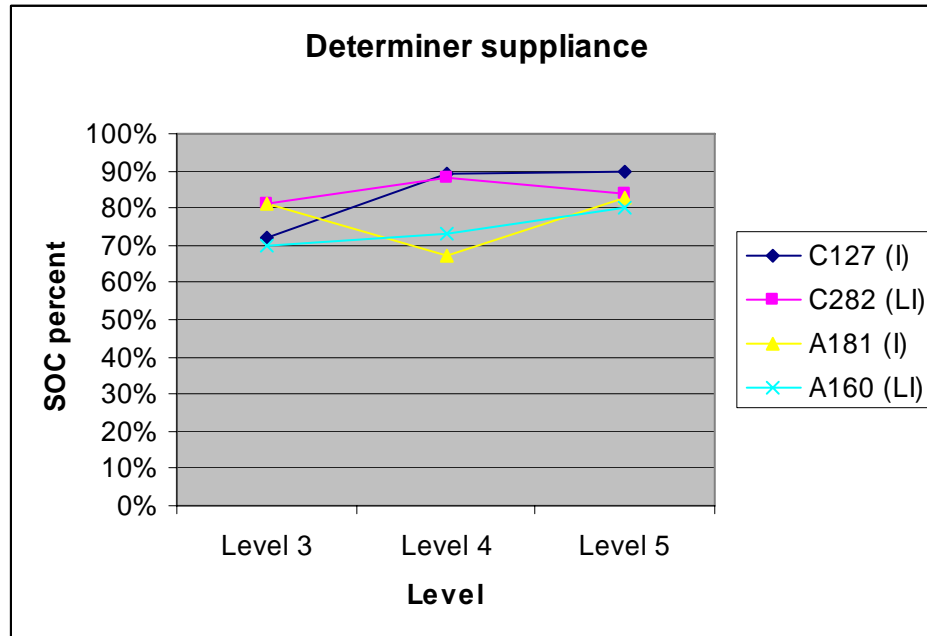


Figure 18. Determiner suppliance by integrative and less integrative learners

Again, rates of suppliance of determiners are relatively high. C127, an integrative learner, is the most successful learner in levels 4 and 5, but the other integrative learner, A181, does not show especially high rates of suppliance. In fact, C282, a less integrative learner, performs as well or better than A181 in all three semesters. Therefore, it must be concluded that suppliance of determiners also fails to distinguish between the integrative and less integrative learners.

Finally, learners' overall error rate was calculated as the number of errors per AS-unit (total errors/number of AS-units). These errors include lexical (word choice) errors, syntactic and morphological errors. The results are presented in Table 23.

Table 23. Errors per clause in highly integrative and less integrative speakers

Learner	Level 3	Level 4	Level 5
C127 (I)	1.4	1.26	.87
C282 (LI)	.76	.67	.73
A181 (I)	.78	1.23	1.0
A160 (LI)	.47	1.2	.97

This information is presented in graph form in Figure 19 below.



Figure 19. Errors per clause in highly integrative and less integrative speakers

The only learner who has a decrease in the number of errors from level 3 to level 5 is C127, the highly integrative Chinese speaker. However, the learner with the fewest errors overall is C282,

the less integrative Chinese speaker. Similarly, the less integrative Arabic speaker (A-160) has fewer errors than the integrative speaker (A-181). These results are not as were expected, based on the Multidimensional Model. Note, however, that the Multidimensional Model does not use overall error counts as part of its predictions.

3.2.4.3 Complexity measures.

Learners' degree of complexity was first calculated as the number of words (omitting repetitions, false starts, and fillers) per AS-unit. That information is presented in Table 24.

Table 24. Words per AS-unit for integrative and less integrative speakers

Learner	Level 3	Level 4	Level 5
C127 (I)	11.2	13.0	9.9
C282 (LI)	11.3	10.3	11.5
A181 (I)	8.4	13.1	14.2
A160 (LI)	8.4	13.6	24.0

The same information appears in graph form in Figure 20:

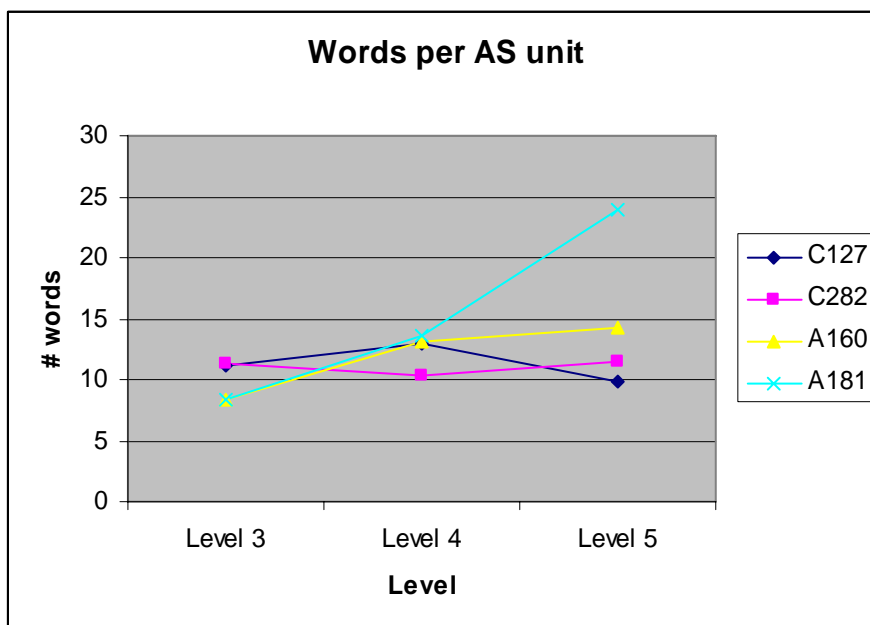


Figure 20. Words per AS-unit for integrative and less integrative learners

Note that the only learner who appears to make significant increases in sentence complexity by this measure is A181, the highly motivated Arabic learner.

Another measure of complexity is the number of complement clauses, relative clauses, and adverb clauses per AS-unit. Those figures are presented in Table 25.

Table 25. Clauses per AS-unit for integrative and less integrative learners

Learner	Level 3	Level 4	Level 5
C127 (I)	.35	.36	.68
C282 (LI)	.39	.46	.31
A181 (I)	.41	.70	.75
A160 (LI)	.18	.53	.48

The same data is presented in graph format in Figure 21.

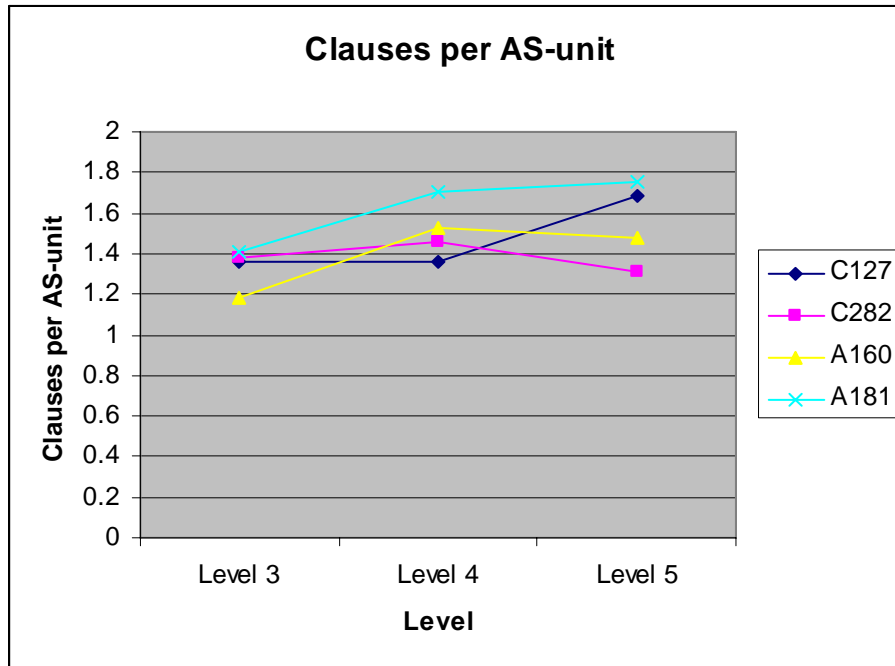


Figure 21. Clauses per AS-unit for integrative and less integrative learners

Figure 21 shows that A181 had the greatest number of clauses per AS-unit. By level 5, the other integrative learner, C127, had the second highest number of clauses per AS-unit.

Therefore, the two complexity measures do show higher scores for the integrative learners.

3.2.5 L1 influence

Based on the analysis of these morphosyntactic features using the emergence criteria, learners of all L1 language groups behave in fairly similar ways. However, there appear to be some patterns of difference between learners of certain L1 groups.

First, the Spanish speakers in this study were more likely to omit subjects than other speakers, regardless of level. Few speakers omitted subjects at all. For those who did omit subjects, exact data on suppliance of subjects in obligatory contexts is provided in Table 26.

Table 26. SOC of subjects for learners who omitted subjects

Speaker, L1, and level	Percent SOC for subjects
Korean 156 Level 2	93 (2 omissions)
Chinese 298 Level 3	97 (1 omission)
Spanish 173 Level 3	93 (6 omissions)
Spanish 130 Level 3	95 (4 omissions)
Spanish 383 Level 4	95 (3 omissions)
Spanish 391 Level 4	90 (7 omissions)

Recall that there are only seven Spanish speakers in the study, meaning that over half of the Spanish speakers omitted subjects, while very few of the other learners did.

A second issue involves the use of relative clauses. The Spanish speakers used them the earliest, starting in level 2. The Arabic speakers appear to have begun using them in level 3. The Chinese speakers also start using them in level 3, although there is more individual variation in this regard. Finally, the Korean speakers appear to use relative clauses beginning only in level 4, with individual variation. The data is displayed in Table 27. A line is drawn to show the pattern of relative clause acquisition.

Table 27. Learners' use of relative clauses

Level	Spanish	Arabic	Chinese	Korean
2	-*_-	N	N	N
	*	N	N	N
		N		N
		-*_-		N
3	N	N	N	N
	-*_-	-*_-	N	N
		-*_-	N	N
		-*_-	*	N
		-*_-	*	
		**		
4	*	-*_-	N	N
	**	*	-*_-	N
		*	*	-*_-
		**	**	*
5	**			

Other issues regarding L1 influence will be covered in the Discussion section.

4.0 DISCUSSION

In the discussion section, each of the research questions posed in section 3.1.6 will be answered with reference to the results.

4.1 ORGANIC GRAMMAR AND THE PATH OF DEVELOPMENT

Can the path described by Organic Grammar/Minimal Trees predict the morphosyntactic development of these learners in these tasks? On the whole, the answer to this question, based on these data, is negative. The following two subsections discuss where the predictions are successful and where they fail.

4.1.1 Emergence of morphosyntactic features

A few predictions of the OG table appear to be borne out. For instance, it is true that for nearly all learners word order, subjects, coordination, and pronouns emerge early, while passive emerges late. Clauses with *because* also generally appear before other types of adverb clauses. However, for the most part, the morphosyntactic development of the learners in this study could not be adequately described by Organic Grammar. Implicational scaling demonstrated that the morphosyntactic elements were not acquired in the predicted developmental order, which

assumes a structure building model in which IP features are acquired before CP features. As a reference, the Organic Grammar predictions are reproduced in Table 28.

Table 28. Organic Grammar (Young-Scholten, Ijuin, & Vainikka 2006, p. 9)

	Word order in declaratives	Types of verbs	Verbal agreement & tense marking	Pronouns	Complex syntax
1	Initially resembles NL	Thematic (main) verbs only	None	Pronouns absent	none
2	Resembles the NL	Thematic verbs; copula 'is' appears	None	Pronoun forms begin to emerge	Formulaic or intonation-based Qs
3	Resembles the TL	Thematic verbs; modals; copula forms beyond 'is'	No agreement, some tense, some aspect, but not productive	More pronouns, but they can still be missing	Qs formulaic or w/o inversion; conjoined clauses
4	Resembles the TL	Thematic verbs, modals, copula forms beyond 'is', range of auxiliaries emerges	Productive tense, aspect, agreement with "be" forms	Pronouns obligatory, 'there' and existential 'it' emerge	Productive Qs, but may still lack inversion; simple subordination
5	Resembles the TL	Complex tense, aspect forms; passives; range of thematic verbs, modals, auxiliaries	Forms usually correct, apart from those newly attempted	Use of 'there' and 'it' beyond stock phrases	All Qs with inversion; complex subordination

All of the learners in this study appear to be at stage 3 or above, based on the presence of pronouns, the copula form *is*, conjoined clauses, and SVO word order for all learners. However, a number of learners fail to produce modals, which is problematic for OG because the lack of modals is a feature of stages 1 and 2. It could be argued that the learners who fail to produce modals, but who have all other stage 3 features, are simply making the transition from stage 2 to 3; however, this cannot be the case, since these learners often produce elements that are predicted to appear in stages 4 or 5: complement clauses, *because*-clauses, and productive past tense marking. Complement clauses are not predicted to emerge until OG stage 5 because they

require the presence of CP. Clauses with *because* are considered to be an element of simple subordination (an element of OG stage 4), and produce productive past tense marking (as measured at 60% suppliance in obligatory contexts) is an IP feature of OG stage 4. None of these elements is predicted to appear in learners who do not have modals.

Past this early stage (level 2 in the ELI; roughly level 3 in OG), it is difficult to categorize learners into stages, because they do not follow the predicted patterns. Learners in OG stage 4 are predicted to have productive tense and aspect (albeit with some errors), simple subordination (*because* clauses and subject relative clauses), and no dropped subjects. Stage 5 learners are predicted to produce complex subordination (complement clauses, adverb clauses, and object relative clauses), productive sentences with expletive subjects, passive, and tense and aspect that is not only productive but also usually correct. However, learners do not fall neatly into these two groups.

The first problem with these predictions is that there is a good deal of variation in individual learners. For example, there are learners (e.g., A338-4) who produce both subject and object relative clauses (stage 5) but not productive sentences with expletive subjects (stage 5); there are also learners (e.g., C118-4) who demonstrate the exact opposite pattern: they produce a large number of sentences with both *it* and *there*,¹⁸ but fail to produce any relative clauses. There are also several students (e.g., K266-2) who use a number of adverb clauses (OG stage 5)

¹⁸ Schachter & Rutherford (1979) found that Chinese-speaking learners of English overused sentences with expletive subjects, especially sentences with *there*. They argue that Chinese speakers are using a discourse strategy based on the Topic/Comment structure in Chinese. This explanation could account for some of the results found here.

but do not produce expletive subjects, relative clauses, complement clauses, aspect, or tense marking.

A second part of the problem with the developmental order predicted by OG is that certain elements are consistently acquired earlier or later than predicted. Let us consider several aspects of the OG table which seem particularly problematic. First, adverb clauses appear to be produced much earlier than is predicted by OG or Minimal Trees. For example, C298-3 produces the following adverb clauses:

(213) These children...maybe steal or make a crime. **When these children can't accept good education**, they can't take a good job. (RSA 3)

(214) Every parents had better have one child. **If everyone follow the rule**, the population in China will maybe decrease (RSA 3)

The student who produced these sentences does not produce relative clauses or tense or aspectual marking. This contradicts OG because adverb clauses are a CP element, while tense and aspect are IP features; additionally, subject relative clauses are predicted to precede adverb clauses. An even more striking case is C271-2, who produces the following two adverb clauses:

(215) **When I was high school student**, I was 14 years old.

(216) **If I have time**, I want to back to Taiwan to see him and tell him my new life.

C271-2 produces no relative clauses, existentials, or past tense marking, and she only produces one modal verb. By the predictions of Organic Grammar, she should not be able to produce stage 5 structures such as these.

Another element that appears out of place in the developmental order is aspect marking, which is particularly late in emerging. OG predicts that it emerges simultaneously with tense, but these data show a different pattern. A large number of learners who show productive tense

(measured at 60% or 80%) do not produce progressive or perfect aspectual marking at all. For instance, K46-3 produces appropriate past tense marking (both regular and irregular), as shown in (217).

(217) The first time I **arrived** there, I **boarded** a kind of ship...I remember that when I **boarded** the ship, I **saw** blue sky and clean lake and also I **could** meet another country friends, so I thought, um. When I **traveled** to Switzerland I **couldn't** speak English very well... (RSA 4)

However, he fails to use any aspectual marking, even when it is required by context: an example is given in (218). Note that either progressive or perfect aspect is required on the verb *change*, but the learner uses simple (presumably default) aspect.

(218) When I **was** a Freshman, there **weren't** many kinds of scholarship in my university. But now, many things **change** in my university. (RSA 3)

Despite using tense properly, this learner produces no aspect-marked forms at all.

In fact, the pattern described above is a common one. The majority of students who are reported as having non-emerged aspect (i.e., with an “N” in the *Asp* column) simply have no tokens of progressive or perfect aspect marking in their production. However, there is another type of aspectual error: that of producing perfect or progressive aspect in inappropriate contexts. For instance, C84-4, who has 82% suppliance of past tense, overuses pluperfect in the following example:

(219) I **had gone** to Hawaii for ten years ago. (RSA 4)

Many other learners overuse progressive, as shown in (220).

- (220) My grandfather of my father is an important person for me because he **was taking** care of us, of all my family in his era and he had a job in the fire station, and he **was earning** a high salary which **is making** a lot of money to build our house
(A45-4, RSA 2)

It is difficult to know what to make of these cases. On the one hand, it could be claimed that these learners have the appropriate morphosyntax in place to produce these features; that is, in order to produce the auxiliary verbs with participles, they must have IP. Therefore, even incorrect uses of aspect marking could be seen as legitimate milestones in an Organic Grammar path of development. However, it is unclear what the learners intend with this marking. Do they intend to mark aspect, tense, or something else entirely? A theory such as the Aspect Hypothesis (Andersen & Shirai, 1994; Bardovi-Harlig, 2000) may help to explain how learners create form/meaning relationships when first acquiring tense and aspect marking. Crucially, the Aspect Hypothesis integrates a multitude of factors in its account of acquisition, including input frequency, L1 background, learning environment, learner characteristics, and especially the inherent lexical aspect in lexical items. These factors may explain why aspect is produced late by the learners in this study. A simple explanation based on functional category (IP vs. CP) apparently does not suffice.

4.1.2 Morphological inflection and developmental order

Recall that an important aspect of Organic Grammar/Minimal Trees (as well as Rapid Profile) is the hypothesis that tense, agreement and aspect marking develop on a predictable path, simultaneously with syntactic IP features. This prediction is not borne out by the data in this study. First, many of the students who produce both simple and complex subordination (CP

features from stage 4 and 5) fail to use past tense marking at even 60% suppliance. In fact, in general there is little pattern between the use of inflection and the emergence of subordination; moreover, this is true for students of all L1s, not just for those whose L1 lacks tense marking. Consider for example these students whose data fail to show a correlation between tense marking and syntactic IP features: S173-3 (Spanish), A182-4 (Arabic), K320-4 (Korean), and C278-4 (Chinese). When the SOC for tense marking is raised to 80%, there is even less of a correspondence between subordination and inflectional marking, which is problematic because students at level 5 are predicted to have nearly all tense and aspect marking correct. These findings are problematic for OG and Minimal Trees.

On the other hand, the findings are consistent with the Missing (Surface) Inflection Hypothesis or “mapping problem” (Haznedar & Schwartz, 1997; Lardiere, 1998a, b; Prévost and White, 2000; White, 2003). The MSIH proposes that second language syntax is dissociated from morphology. Proponents of this theory would argue that morphological marking cannot be included in a developmental timetable, as various factors specific to the morphological component may inhibit learners’ morphological production.

For instance, Lardiere (1998a) examines the spontaneous oral production in English of “Patty”, a native Chinese speaker who has lived in the United States for many years. Lardiere notes that although Patty frequently uses bare verbs, uninflected for tense, her use of nominative and accusative case on pronouns is completely appropriate. The conclusion is that Patty must have an intact Tense feature in her grammar, since according to current Generative theory, Tense licenses nominative case in subjects. However, it is not expected that a learner with Tense should have difficulty realizing overt verbal morphology. Lardiere reasons that Patty’s difficulty lies not in the syntax, but rather closer to the “surface”, that is, in the morphophonological

component. Using the framework of Distributed Morphology (Halle & Marantz, 1993), Lardiere concludes that a breakdown occurs in L2 speakers' mapping component, that is, the linguistic element that matches vocabulary items with feature bundles in the syntax. Further evidence for this account is provided by Goad, White & Steele (2003). They examined production data from Chinese speaking learners of English and found that many of the speakers omitted past tense marking significantly more often in certain phonological contexts, specifically when past tense marking could not be incorporated into the phonological word in accordance with Mandarin phonological rules. They conclude that the speakers have not yet acquired English phonological rules, which prevents them from mapping appropriate vocabulary forms to feature bundles in the syntax (but cf. Snape, 2006).

Supporters of the Missing Surface Inflection Hypothesis might argue that the morphological component of the Organic Grammar placement test should be removed. Indeed, it was found that removing morphological measures from the implicational table (in Table 14) did improve the reproducibility and scalability of the table. The fact that the improvement was not large is most likely due to the large number of "gaps" in data for morphological marking. It is predicted that if a more thorough account of learners' morphological marking is included in such a table (perhaps through the use of targeted elicitation techniques), then the improvement in scalability and reproducibility will be more obvious. If this is true, then morphological marking should be removed from a developmental scale.

It is important to note, however, that removing tense, agreement and aspect marking from a placement or proficiency measure may seem counterintuitive, or at least undesirable from the point of view of language testers or language program administrators. Morphological marking may be very important to both instructors and learners; tense and aspect are often crucial to clear

communication, while errors in agreement marking may quickly tag a learner as “foreign” or “uneducated”. In fact, tense and aspect marking are the only two grammatical forms that are specifically referred to in the ACTFL proficiency guidelines. Therefore, even if tense, agreement and aspect marking are subject to acquisitional principles that are different from other syntactic elements, it may still be valuable to include them in evaluations of student proficiency. However, keeping measures of morphological marking separate from measures of syntax may be more appropriate.

4.2 RAPID PROFILE AND THE PATH OF DEVELOPMENT

Question 2. Can the path described by Rapid Profile account for the morphosyntactic development of these learners in these tasks? To some extent.

An implicational table indicated that the Rapid Profile measure was both reproducible and scalable. That is, in the learners in this study, word order, plural, regular past tense marking, and possessive pronouns emerge before sentences beginning with a non-subject, object pronouns, 3rd person singular marking, and *wh* noun phrases. The use of *-ing* forms is the only element that is not reproducible in this table, probably because it is associated with aspect, which emerges in a less predictable way than other elements.

Unfortunately, important Rapid Profile elements could not be included in the current study because the data are the product of spontaneous production rather than elicitation of specific structures. Students in this study did not produce questions (or phrasal verbs or possessive marking on nouns); without this data it is difficult to make a firm assessment of the

validity of the Rapid Profile measure with spontaneous production data. In and of itself, however, the methodological technique used here should not affect the order of emergence; Pienemann (1998) predicts that the order predicted by Processability Theory is consistent across data types. The current issue is whether the features produced in spontaneous production can be used in a placement test.

Keeping this in mind, let us consider the results regarding the Rapid Profile implicational table in this study. The table was found to be reproducible and scalable, which indicates a predictable developmental path for these learners. However, that developmental path may not provide much usable information for placement or proficiency testing. The RP table suffers from the same problem as does the OG table, in that a large portion of it belongs to the “emerged” side. That is, there are not enough elements to distinguish between intermediate and advanced learners, because nearly all the students have emerged word order, plural, past tense marking, possessive pronouns, and sentences beginning with a non-subject element. In other words, only a few elements actually do any “work” in the table as indicators of development. In fact, only the *Cancel Aux 2nd* rule is not emerged in most learners. Because this rule may not be as likely to be used in spontaneous production as many other forms (such as modals), even this measure may not be useful in determining which learners are the most advanced. Perhaps learners simply use this structure infrequently, and its absence in the data has little to do with development. This possibility could partially explain the fact that learners in levels 3, 4 and 5 are scattered throughout the table.

A related problem was pointed out by Hudson (1993), who argued that Processability and Rapid Profile are only useful when measuring a very short stage of early development. He

argues that the five Processability stages¹⁹ only cover the first two and a half levels of seven levels of language proficiency (i.e., 36%),²⁰ leaving most of the path of L2 change undiscussed and unanalyzed. The same problem exists in these data, where only learners at the beginning levels (ELI level 2) can be distinguished through use of RP. ELI levels 3 and higher have the various elements to differing degrees.

This finding may be due to the missing elements in the table. However, an additional possibility is that the problem is within Rapid Profile itself. Notice that, even when questions are included in the profile, there are a large number of morphosyntactic elements that are simply left out: modals, relative clauses, adverb clauses, various kinds of complement clauses, other types of inflection (such as past participle forms), expletives, articles, do-support, auxiliaries, copula, and passive.

Because few of the elements of Rapid Profile were included in this study, little more can be said about the validity of RP itself. However, it may be possible to include a few of the measures in a placement test based on spontaneous production data.

¹⁹ Hudson declines to include the 6th level of development, claiming that it was not supported by empirical evidence.

²⁰ Hudson refers to seven levels of proficiency based on the proficiency scale of the Adult Migrant Education Service (AMES) Scale.

4.3 A DIFFERENT PATH OF DEVELOPMENT

Question 3: If the paths described by RP and OG do not account for this data, can the morphosyntactic elements appear in a different, but still predictable, order? Yes, to some extent.

In the Results section, the Organic Grammar implicational table was revised to produce the most reproducible and scalable table possible, without removing any elements. In that revised table, the elements appeared in the following order, starting with the earliest emerged. Items that appear on the same line can change order without affecting the table.

1. Word order/Pronouns/Subjects
2. Copula forms beyond 'is'/Coordination of clauses
3. Complement clauses
4. Past tense marking as measured by 60% suppliance in obligatory contexts
5. *Because*-clauses
6. Agreement as measured by 60% suppliance of 3rd person singular *-s* in obligatory contexts
7. Modals
8. Adverb Clauses (except those with *because*)
9. Do-support
10. Use of expletive subjects
11. Relative clauses
12. Aspect (use of perfect and/or progressive)
13. Passive

In this order, the implicational table is both reproducible and scalable. However, it is based only on measures from Organic Grammar. In order to create the most complete table, measures were combined from OG and RP into a single implicational table, which was then maximized for

reproducibility and scalability. Some of the measures with high error rates were eliminated to create a better table.

Unfortunately, it was not possible to use all of the measures in RP and OG to create a table with a coefficient of reproducibility over .9 and a coefficient of scalability over .6. That is, when the morphosyntactic elements from both systems are combined, the table inevitably has a high number of errors. This result is probably due to the fact that neither system is detailed enough to make accurate predictions for learners from levels 3 to 4. Remember that limited rearrangement of learners is possible in the implicational table; learners appear in the table in the order that creates the fewest number of errors. Therefore, the order of learners from top to bottom of the table is to some extent due to the rearrangement of learners to create the best table, rather than a real order from least to most advanced. Table 29 was created as the table with the best predictions for all levels, using the measures of Rapid Profile and Organic Grammar.

Table 29. Implicational table combining measures of RP and OG for maximum scalability

Student	Cop	Co.	Cmp	B/c	Adv			Do/s	RC	Asp	Pv	Cncl	
					1st	Mod	AC					Aux	
K123-2	*	*	*	N	N	N	N	N	N	N	N	N	N
A148-2	_*	*	N	*	N	N	N	N	N	N	N	N	N
A25-2	*	_*	_*	*	N	N	N	N	N	N	N	N	N
A12-2	N	*	_*	N	*	N	_*	N	N	N	N	N	N
K267-2	_*	*	_*	N	*	N	N	**	N	N	N	N	N
C126-2	_*	_*	_*	_*	*	N	N	N	N	N	N	N	N
K269-2	*	*	N	*	*	N	_*	N	N	_*	N	N	N
K266-2	**	_*	N	_*	N	_*	_*	*	N	N	N	N	N
K156-2	**	_*	_*	N	*	_*	_*	N	N	N	N	N	N
K167-2	**	*	*	*	*	_*	_*	N	N	N	N	N	N
C271-2	**	*	*	_*	*	_*	*	N	N	N	N	N	N
C127-3	**	*	**	*	*	*	N	N	N	N	N	N	N
K111-3	**	*	_*	*	*	*	*	N	N	N	N	N	N
C118-4	**	N	**	*	*	**	*	N	N	N	_*	N	N
S366-2	*	_*	**	_*	*	*	N	N	*	N	N	N	N
A45-2	*	*	_*	*	*	_*	N	N	_*	N	N	N	N
A170-2	**	*	N	_*	*	**	*	N	_*	N	N	N	N
C298-3	*	*	*	N	*	**	*	**	N	N	N	N	N
K300-3	*	*	_*	*	*	_*	*	_*	N	N	N	N	N
A338-4	**	*	**	*	*	**	*	**	N	N	N	N	N
A159-3	*	*	_*	_*	*	*	_*	_*	N	_*	N	N	N
S173-3	**	*	*	*	*	*	**	_*	N	_*	N	N	N
A182-4	*	*	**	N	*	*	**	_*	N	*	_*	N	N
S130-3	**	*	**	*	*	**	_*	N	_*	*	N	N	N
K217-4	**	*	**	N	*	**	**	*	N	N	N	*	N
K101-3	**	*	*	*	*	**	*	_*	N	N	N	N	N
K46-3	**	*	*	_*	*	N	*	*	N	_*	N	N	N
C177-3	**	*	**	*	*	*	*	_*	N	N	N	N	N
S362-2	*	*	**	*	*	**	_*	N	_*	N	N	N	N
A279-4	**	*	**	*	*	**	*	_*	*	N	N	N	N
C84-4	*	*	**	*	*	**	**	_*	**	N	N	N	N
A157-3	**	*	*	*	*	*	_*	N	**	N	N	*	N
C301-3	**	*	**	*	*	**	**	**	*	N	N	N	N
C282-3	**	*	**	*	*	_*	_*	*	*	N	_*	N	N
K276-4	**	*	**	*	*	**	**	*	N	_*	_*	N	N
A160-3	*	*	_*	*	*	*	*	*	_*	_*	N	N	N
T397-5	**	*	**	*	*	*	N	N	*	_*	_*	N	N
A199-4	**	*	**	*	*	_*	**	_*	_*	*	N	N	N
C201-4	**	*	**	*	*	*	*	_*	_*	*	N	N	N
S391-4	**	*	**	*	*	**	*	*	**	*	N	*	N
S383-4	**	*	**	*	*	**	_*	N	*	**	N	*	N
K320-4	**	*	**	*	*	**	**	**	*	*	N	N	N
A163-3	**	*	**	*	*	_*	*	*	_*	**	N	N	N
A161-3	**	*	**	*	*	*	**	_*	_*	*	N	N	N
K275-4	*	*	**	_*	*	*	**	*	_*	_*	N	*	N
S100-5	**	*	**	*	*	**	**	_*	**	_*	**	N	N
J274-5	**	*	**	*	*	**	**	_*	**	N	_*	*	N
C278-4	*	*	**	*	*	**	**	*	*	_*	_*	*	N

For this table, the Coefficient of Reproducibility is calculated below.

$$(221) \text{ C of R} = 1 - (51/576) = .91$$

This figure indicates a reproducible table. The Coefficient of Scalability is calculated in (222):

$$(222) \text{ C of S} = .21/(1 - .7) = .7$$

A C of S of a minimum of .6 to .65 is considered to indicate a scalable table, so we can conclude that this is a real progression.

The question remains, however: how can it be explained as a developmental path? It does not follow the predictions of Minimal Trees or Processability Theory. It also suffers from the same problems as the implicational tables based solely on OG and RP. First of all, its ordering of students is quite different than the order produced through the traditional placement tests used at the Pittsburgh ELI. Note that if the developmental measures produced the same result as the traditional placement measures, the level 2 students would be followed by the level 3 students, who would be followed by the level 4 students, etc. (i.e., 22233334445). However, only the level 2 students are consistently ranked at the top of the chart (i.e., the lowest developmental stage); students of levels 3 and 4 are mixed throughout the rest of the chart, with only a tendency for level 4 students to appear lower in the chart. The fact that the order of students deviates so strongly from the order predicted by traditional placement measures indicates that it will not suffice as a placement measure in and of itself. However, the elements still may be useful as contributions to a proficiency measure that includes other measures.

A second problem with this table, as well as the others, is that it provides a more meaningful distinction for students in early levels. Many students at ELI levels 5, 4 and even 3 produce nearly all of the morphosyntactic elements in the table. Yet these students clearly have

not reached the highest proficiency possible. Consider these excerpts from the speech of C282-3, a Chinese speaker who produces all but two of the morphosyntactic elements in the chart.

(223) We treat the pets like only the animal. For example, dog is only guides the house.

The pets they to they to catch cats, and cow they have to work with farmer, but...

(RSA 1)

(224) President's wife and president's daughter, they spent money illegal. They used

the money which don't possess themselves. But the fact was found. They said

many excuses that's very funny and unreasonable. (RSA 3)

Learners at the highest point of development (as measured by OG and RP) make a large number of errors: syntactic, morphological, and lexical/idiomatic; for instance, prepositions are omitted; agreement endings are incorrect; tense marking is misused; word order in phrases is incorrect, etc. However, these errors are irrelevant in these developmental tables because emergence is used as a measure rather than accuracy. It may be possible to expand the reach of the table by including error data or information about the number of times an element is used.

4.4 FIRST LANGUAGE INFLUENCE

Question 4. To what extent does a learner's first language affect the path of morphosyntactic development? It only affects development in a few areas.

Organic Grammar and Rapid Profile both predict a minimal influence of the native language on the path of second language morphosyntactic development. The influence of the L1 is an important issue, because if learners with different L1s had different paths of morphosyntactic development, then there would be no way to use a single developmental scale as a measure of all

learners. By extension, we can conclude that it would also be impossible to create a single morphosyntactic proficiency measure for all learners.

It appears, however, that Spanish, Korean, Chinese and Arabic speakers produce the morphosyntactic elements in this study in a similar order (at least, to the extent to which an order can be predicted). That is, there are very few areas in which one native language group appears to have greater or lesser success, by these measures. Additionally, the L1s do not cluster in the implicational tables, which would suggest similar performance or ability on a certain element; rather, the learners are scattered throughout the table.

There are, however, a few areas in which L1 seems to make a difference. The most striking of these is the suppliance of subjects. It was found that virtually all the learners supplied nearly 100% of required subjects.²¹ There were six exceptions, however. Interestingly, these exceptions were not all at beginning levels, either as measured by ELI level or by OG stage. Crucially, four out of six of these learners were Spanish speakers, which means that a majority (4 out of 7) of the Spanish speakers omitted subjects. The table with details about subject use is reproduced in Table 30.

²¹ Unfortunately, by this measure, nearly all of the learners, including ELI level 2 students, are placed into OG stage 4, surely not a desired result.

Table 30. Learners' use of subjects

Speaker, L1, and level	Percent SOC for subjects
Korean 156 Level 2	93 (2 omissions)
Chinese 298 Level 3	97 (1 omission)
Spanish 173 Level 3	93 (6 omissions)
Spanish 130 Level 3	95 (4 omissions)
Spanish 383 Level 4	95 (3 omissions)
Spanish 391 Level 4	90 (7 omissions)

Examples of dropped subjects from the learners are provided below.

- (225) My Mother's side still living there. Is a kind welcome home safe port that is always there for you, so is what I like to do. (S391-4, RSA 2)
- (226) When I will go to the beach, always play volleyball and tennis with my sisters and friends. Sometimes in my free time go to walk in the park. (S130-3, RSA 1)
- (227) Her hair is long and straight. Is very pretty, is brown. (S173-3, RSA 1)

These learners do not drop all subjects, or even all subjects that could be grammatically omitted in Spanish. They simply have the tendency to omit subjects more than speakers from other backgrounds. Note that Spanish speakers in ELI levels 3 and 4 (and at high stages of OG development) are still omitting subjects. This fact is not predicted by OG, which predicts that learners will eliminate null subjects when they acquire IP.

The data also stand in contradiction to the predictions of Rapid Profile. In the Rapid Profile system, suppliance of subjects is based on degree of integratedness, or motivation and desire to assimilate into the target culture. It is possible that these Spanish learners are less

integrated than the others, but if that is the case, then they should also display lower performance on other variational features, such as accuracy of marking of morphological endings. This is not the case. In fact, these learners are highly successful and appear quite advanced in terms of their production of morphosyntactic elements.

Based on these observations, it may be logical to conclude that Spanish speakers are omitting subjects because of a “transfer effect” from Spanish. Recall that Spanish is a pro-drop language which allows the omission of pronoun subjects when they are clear from context. Learners may occasionally be allowing pro-drop to operate in their Spanish/English interlanguage. However, recall that Arabic, Chinese and Korean *all* allow the omission of subjects (as well as objects, in some circumstances). This fact makes it more difficult to explain why Spanish learners appear particularly likely to omit subjects. One possibility is that the licensing of null anaphora operates differently in Spanish. Several proposals have suggested that null anaphora in topic-prominent languages such as Chinese and Korean are licensed differently than in languages such as Spanish (e.g., Huang, 1989). It may also be the case that Arabic has characteristics of topic-prominent languages, as argued by several researchers (e.g., Doron & Heycock (1999), Alexopoulou, Doron, & Heycock, (2001)). Spanish is clearly a subject-prominent language; perhaps this characteristic allows the Spanish-speaking learners to license optional null subjects in their Spanish-English interlanguage.

Another area where L1 seems to have influence is relative clauses. Based on these data, relative clauses appear to emerge earliest in Spanish speakers, then in Arabic speakers, then in Chinese speakers, and finally in Korean speakers. This information is given in Table 31.

Table 31. Learners' use of relative clauses

Level	Spanish	Arabic	Chinese	Korean
2	-*_-	N	N	N
	*	N	N	N
		N		N
		-*_-		N
				N
3	N	N	N	N
	-*_-	-*_-	N	N
		-*_-	N	N
		-*_-	*	N
		-*_-	*	
4	*	-*_-	N	N
	**	*	-*_-	N
		*	*	-*_-
		**	**	*
5	**			

Remember that Spanish and Arabic relative clauses are similar to English relative clauses. First of all, Spanish, Chinese and Arabic all use free relativizers to introduce clauses (which can be omitted in Arabic, but not in Spanish). However, Korean uses verbal suffixes to indicate relativization. It seems logical to conclude that the presence of free relativizers in Arabic, Chinese and Spanish may have helped these speakers to identify relativizers in English and therefore to find relative clauses in input. Second, Arabic and Spanish more resemble English in that they have postnominal relative clauses, while Chinese and Korean have pronominal relative clauses. Finally, Comrie (1989; 2007) has also suggested that the structure of relative clauses in Korean and Chinese is radically different from that in languages such as English and Spanish, which may also account for the differences found here.

There were no other areas in which L1 influence appeared to create a significant effect on emergence. It is important to note, though, that an emergence measure leaves out detail about the use of the elements. For instance, a learner is considered to have emerged modals after a single productive appearance of a modal + verb, but the exact choice of modal does not affect the outcome. Errors occurring before or after the emergence of modals (i.e., accuracy) are also not included in the measure. Therefore, it is possible that some morphosyntactic elements are used differently by learners of different L1s, but this information is not examined here.

4.5 INFLUENCE OF MOTIVATION AND EXPOSURE TO ENGLISH

Question 5. To what extent does motivation, attitude, and exposure to English affect the path of morphosyntactic development or the accuracy of marking on morphosyntactic elements? It may have some effect on complexity, but the influence is minor.

Organic Grammar predicts that the path of development will be the same for all learners, regardless of their motivation or amount of exposure to English. Some learners may move more quickly through the stages, but the path of development remains the same. Indeed, this assertion appears to be generally correct, based on the four learners whose data were examined. They all appeared to progress through the stages in the same order and in a similar manner.

The Multidimensional Model of Meisel, Clahsen & Pienemann (1981) and Clahsen, Meisel & Pienemann, as well as Pienemann (1998) predicts that learners with higher integrativeness will have greater success on variational features, including suppliance of subjects, copulas, and determiners. However, this was not found to be the case for the four

learners whose data were examined. All the learners produced high suppliance of subjects and copulas, and the suppliance of determiners did not seem to depend on integrativeness. These results confirm findings by Schmidt (1983), who found that a strong degree of integrativeness and acculturation does not necessarily lead to grammatical gains. Schmidt examined the progress of an L2 learner of English in Hawaii. Even though “Wes” was highly interested in American/Hawaiian culture and needed to work there, his grammatical development was very limited.

The only difference that was found between the integrated and the less integrated learners was in terms of their linguistic complexity. These results suggested that more integrated learners might produce longer, more complex sentences. There are several possible explanations for these results. First, it is possible that motivated learners with large amounts of contact with native speakers are exposed to more sophisticated language use, which they are then able to integrate into their own speech patterns. This explanation is in the spirit of the Multidimensional Model. A more mundane but equally reasonable explanation is that the more integrative learners were simply more motivated to receive a higher grade on the RSA and therefore attempted to deliver a more sophisticated speech.

Before concluding that the Multidimensional Model is incorrect, it is important to consider a few limitations in the data. First, it is possible that the effect was not found here due to the homogeneity of the participants. That is, there may not enough variation in the integrativeness of the learners to produce a significant difference in results on any of the measures. Second, it is possible that these learners have already attained a much higher degree of proficiency than the learners who are generally profiled in the Multidimensional Model. That

is, most of these learners have progressed quite a long way on the developmental path, which would not be possible unless they had avoided or moved past certain behaviors.

At any rate, the results presented here must be considered preliminary. Further research is required to make any definitive statements on the subject.

4.6 “COMPLEXITY” IN THE PATH OF DEVELOPMENT

In Minimal Trees theory, Vainikka & Young-Scholten (1994; 1996a, b; 1998a,b) have predicted that the initial state of second language learners is a VP transferred from their native languages. Learners must then acquire a Functional Phrase (FP), Inflectional Phrase (AgrP or IP) and finally Complementizer Phrase (CP). MT therefore predicts that elements associated with IP should emerge before those associated with CP, ideally all at once in a “cluster.” Repeated here for convenience is a list of elements associated with each functional phrase.

(225) **Elements associated with IP:** subjects, auxiliaries, do-support (as in negation), modals, tense, agreement and aspect marking, expletive subjects

Elements associated with CP: adverb (adjunct) clauses, wh-questions, noun (complement) clauses, infinitival complements of verb (*want to eat, like to sleep, etc.*), relative clauses

Interestingly, the predictions of Organic Grammar do not directly follow those of Minimal Trees. For instance, expletive subjects are not associated with CP, and yet they appear in the highest (i.e., last-emerging) stage of OG. Similarly, both subject and object relative clauses require CP, but the two types of clause are separated in OG, with subject clauses belonging to “simple subordination” (OG stage 4) and object clauses belonging to “complex subordination” (OG stage

5). Finally, OG places adverb clauses with *because* in the category of simple subordination, despite the fact that these clauses are generally considered to require CP.

These examples demonstrate that the creators of the Minimal Trees hypothesis tacitly recognize more than one layer of complexity in morphosyntactic structure, even though it is not allowed by their theory. That is, forms and structures may be difficult for reasons that are far more nuanced than simply which functional category they are associated with. Several factors other than functional category may help to explain the order of emergence that was found for the learners in this study. I will discuss two here: semantic complexity and syntactic gaps (displacement).

First of all, semantic issues may make it difficult to use certain forms. To give an example from this study, it is far too simplistic to claim that aspect should be acquired along with tense and agreement marking because all are associated with Inflectional Phrase. As I have already mentioned, the Aspect Hypothesis (e.g., Andersen & Shirai, 1994) argues that input frequency, learning situation, learner characteristics, and inherent lexical aspect all contribute to the creation of form/meaning relationships by early learners. Let us consider some of these challenges in acquiring aspect in English.

First, the proper use of English aspect requires the mastery of subtle semantic distinctions that can be very difficult to articulate. For instance, the meaning of present progressive changes depending on the inherent lexical aspect of the verb or verb phrase. Consider the following two examples:

(226) The plane is landing.

(227) She is coughing.

In (226), the action of landing the plane is in progress. This sentence may be uttered when the plane begins its descent, when the wheels are lowered, or when the plane is on the runway; it cannot, however, grammatically be uttered after the plane has completed a landing. Now consider (227). The interpretation here is iterative; that is, the coughing occurs not once but multiple times. The difference in meaning is due to the fact that *land* is an achievement verb, while *cough* is a semelfactive verb (Smith, 1997). Learners must be able to make these subtle distinctions if they are to use aspect correctly.

To make the situation more difficult, native speakers may use aspect inconsistently. For instance, native English speakers are notoriously inconsistent about using the past perfect. For many speakers, the two sentences presented in (228) and (229) are interchangeable.

(228) I had to go home because I left the oven on.

(229) I had to go home because I had left the oven on.

To give another example, for many native English speakers, there is little difference between simple and perfect aspect in (230) and (231):

(230) Did you eat yet?

(231) Have you eaten yet?

Learners of English must come to understand when the distinction between aspects is semantically important and when it is not.

Finally, the use of aspect by native speakers may contradict the instruction that learners receive from textbooks or teachers. For example, students are typically taught that stative verbs such as *know*, *love* and *own* cannot be used in the progressive. However, it does not take very long for learners who are living in the United States to be confronted with the following slogan:

(232) I'm loving it.

The point to take from these examples is that it is logical to expect aspect to emerge differently than tense or agreement. The reasons may have little to do with syntactic structure, but rather be due to semantic complexity and seemingly contradictory input.

Another factor that may influence the emergence of a structure is whether it has gaps. A gap is created when there is movement, in traditional transformational grammar. For instance, consider the passive construction in (233):

(233) The ball_i was kicked t_i.

Traditionally, this structure is created by moving the direct object *ball* into subject position (Specifier of IP). A gap is left after *kick*, where the direct object would normally appear.

Another case of movement is given in (234):

(234) I don't know what_i you like t_i.

In (234), the pronoun *what* has been displaced from its position after the verb, leaving a gap after the verb *like*. Recall that Clahsen & Felser (2006) claim that it is precisely empty categories and gaps that are lacking in the “shallow” processing of L2 learners.

Notice that both of these constructions—passive and *wh*- noun clauses—appear late in learners' development. The late appearance of both of these structures is highly reproducible (i.e., predictable), as measured by the coefficients of reproducibility in the implicational tables. However, both OG and RP fail to use this information completely. OG does not refer specifically to *wh*- noun clauses, and simply includes them with elements such as noun complements with *that*, as shown in (235).

(235) I think **that she is nice**.

The two types of complements are considered equivalent because both require CP. On the other hand, RP does not include the use of passive at all, presumably because it does not involve the relation of features in the same way that morphological inflection or questions do.

Because neither OG nor RP allows the inclusion of semantic or structural issues such as these in the developmental scale, neither one may be able to account for the use of certain forms and structures. That is, the scales may be too simplistic to predict the development that they are designed to predict.

Let us consider one final issue regarding complexity that is missed by both OG and RP, that is, variation within a single morphosyntactic element type. For instance, in OG it is generally assumed that if a learner can produce one kind of (non-chunked) modal verb, she can produce them all, just as in RP it is assumed that producing a single (non-chunked) regular plural form indicates that processing of all regular plurals has emerged. However, all forms were not produced equally in the data. First, consider modals. *Can* appeared most often, followed by *will* (to indicate future). Learners also used *could*, *would*, *should* and *might*, but much less often. A count of learners' use of modals appears in Table 32.

Table 32. Learners' use of modals in spontaneous production data

Can	Will	All other modals
162	98	51

Generally speaking, if learners produced *might*, *should*, *would* or *could*, they also produced *can* and *will*, although the reverse is not true. It is possible, therefore, that there is an implicational order of modal acquisition, within the larger implicational table.

Adverb clauses also appear to emerge in a particular order. Most learners begin by using *because* clauses. The next clauses that they produce are generally *when* and *if* clauses and possibly *before* or *after* clauses. Clauses with *until* and *although* or *even though* appear only at upper levels (ELI 4 and 5). This is not a strict implicational order, but rather a strong tendency. The percentages of adverb clauses that learners produced in this data are presented in Table 33²².

Table 33. Learners' use of adverb clauses in spontaneous production data

Because	When	If	Before	After	Since	Although/ Even though	Until
226	133	82	12	11	4	2	2

There are a number of possible explanations for the order of emergence of modals and adverb clauses. The modals and clauses that appear first may be simpler for semantic reasons, or they may be more prevalent in the input, including in instructional texts. On the other hand, the order of emergence seen here could be a task effect. That is, it may be that the types of questions that are posed to these learners elicit certain types of verbs and clauses more readily than others. It would be useful to investigate this issue further with a different method of data elicitation and analysis. The precise reason for these of emergence lies beyond the scope of this project, but clearly the order of emergence cannot be explained by referring only to type of functional category (CP or IP).

²² Note that frequency of use does not necessarily correlate with difficulty or complexity. The tables are presented as a point of reference only.

An additional factor not dealt with by either OG or RP is that of quantity of morphosyntactic elements. The combined implicational table (Table 29) shows that in general, learners at more advanced levels produce not only a greater variety of morphosyntactic elements, but also a greater number of each type. That is, when an element first emerges, it may be produced very sporadically, but as time passes, learners use the element more often, partially because they are able to produce more speech in a shorter period of time. Therefore, it may be useful to consider the number of tokens of morphosyntactic elements along with the type and variety of morphosyntactic elements when analyzing learner data in terms of proficiency level.

5.0 CONCLUSION

5.1 DEVELOPMENTAL FEATURES AND PLACEMENT TESTS

I have presented a number of problems with using the current morphosyntactic development measures as a proficiency/placement test. However, this study has shown that a number of morphosyntactic features do emerge in a predictable order, and there may be important contributions that can be made to current proficiency/placement features based on this information. Primarily, this information could serve to supplement the proficiency and placement tests that are currently available by providing concrete morphosyntactic “benchmarks” for learners as they progress. These benchmarks could provide helpful, objective information to interviewers who are attempting to judge a learner’s level.

For example, the current ACTFL rubric could be improved by the inclusion of concrete morphosyntactic elements. The current version is problematic because it uses a large number of relative terms. For instance, consider the following descriptors taken from the ACTFL scale. Quotes are from ACTFL, 1999; italics are mine.

- Novice High speakers can *generally* be understood by *sympathetic* interlocutors used to non-natives.
- Intermediate Low speakers can *generally* be understood by *sympathetic* interlocutors, particularly by those accustomed to dealing with non-natives.

- Intermediate Mid speakers are *generally* understood by *sympathetic* interlocutors accustomed to dealing with non-natives.
- Intermediate High speakers are *generally* understood by *native speakers* unaccustomed to dealing with non-natives.
- Advanced Low speakers are *readily* by *native speakers* unaccustomed to dealing with non-natives.

Notice that the descriptors differ from one another in very subtle ways. Novice High through Intermediate Mid learners are thought to be understandable by sympathetic interlocutors, while Intermediate High and Advanced Low learners can be understood by “native speakers.” It is of course difficult to define who is sympathetic and who is not. Similarly, it may be difficult to define who is accustomed to dealing with non-natives and who is not. (Presumably, of course, a student undergoing an Oral Proficiency Interview hopes that an examiner will be both sympathetic and accustomed to dealing with non-natives.) Even more problematic is the fact that the only difference between the descriptors for Intermediate High speakers and Advanced low speakers is the adverb: *generally* versus *readily*.

Several other descriptors are equally vague or relative. Consider the following extracts (ACTFL 1999; italics are mine).

- Speakers at the Novice High level are able to manage successfully *a number of uncomplicated* communicative tasks in *straightforward* social situations.
- Speakers at the Intermediate Low level are able to handle successfully *a limited number of uncomplicated* communicative tasks by creating with the language in *straightforward* social situations.
- Speakers at the Intermediate Mid level are able to handle successfully *a variety of uncomplicated* communicative tasks in *straightforward* social situations.
- Speakers at the Intermediate High level are able to handle successfully *many uncomplicated* tasks and social situations.

- Speakers at the Advanced Low level are able to handle *a variety* of communicative tasks, although somewhat haltingly at times.
- Speakers at the Advanced Mid level are able to handle with ease and confidence *a large number* of communicative tasks.

Different interviewers will of course have very different ideas about what constitutes an uncomplicated communicative task or a straightforward social situation. Notice that here again, relative terms are used. The learner may be able to handle *a limited number, a number, a variety, many, or a large number* of communicative tasks. There is no way to precisely define these terms.

The point is not that the ACTFL guidelines serve no purpose. Measuring communicative competence is, as discussed in section 2.1, very difficult, and there are no perfect solutions yet. However, by making reference to specific grammatical elements, the scale could be “anchored” to real-world, objective benchmarks of proficiency. Consider the excerpts from the scale that are presented in Table 34. These excerpts are the descriptors that refer to morphosyntax, or that could be enhanced with concrete morphosyntactic terms. Notice that only tense and aspect are the only grammatical elements that are directly mentioned. (In this chart, the most basic learners are at the top, while the advanced learners appear at the bottom.)

Table 34. Selected descriptions from ACTFL scale (1999)

ACTFL level	Description
Novice Low	<ul style="list-style-type: none"> • formulas only
Novice Mid	<ul style="list-style-type: none"> • formulas + words & phrases
Novice High	<ul style="list-style-type: none"> • syntax strongly influenced by L1 • can produce expansions of learned material and stock phrases
Intermed. Low	<ul style="list-style-type: none"> • syntax is strongly influenced by L1 • provides short statements assembled from pieces of memorized formulas or from statements used by interlocutor • can ask a few questions • can accomplish uncomplicated communicative tasks
Intermed. Mid	<ul style="list-style-type: none"> • can produce a variety of questions • have difficulty linking ideas • have difficulty manipulating time and aspect • have inaccuracies in grammar/syntax
Intermed. High	<ul style="list-style-type: none"> • can exchange basic information • can narrate and describe in major time frames • breakdowns possible in maintaining narration or description in appropriate major time frame • errors may be evident
Advanced Low	<ul style="list-style-type: none"> • can narrate and describe in all major time frames • control of aspect lacking at times • “grammatical roughness”
Advanced Mid	<ul style="list-style-type: none"> • have good control of aspect • have ability to narrate and describe in all major time frames • lack the ability to consistently provide a structured argument • can narrate, describe, explain or tell anecdotes
Advanced High	<ul style="list-style-type: none"> • can explain in detail • can narrate fully and accurately in all time frames • can construct hypotheses, but patterns of error appear • may speak abstractly, but prefer to speak concretely • may need to use description or narration in place of argument or hypothesis
Superior	<ul style="list-style-type: none"> • can separate main ideas from supporting information • can explain opinions • can provide structured argument to support opinions. • can construct and develop hypotheses to explore alternative possibilities • can provide abstract elaborations

Some of the morphosyntactic markers in the ACTFL table disagree with the findings in this study. For instance, I found that some learners of advanced levels of proficiency still failed to realize tense and agreement consistently, but the ACTFL scale assumes that learners will successfully mark tense at relatively early stages (Intermediate High/Advanced Low). Another problem is that the ACTFL scale makes frequent reference to accuracy, while the systems investigated in this study refer only to emergence.

However, there are many points of compatibility between the ACTFL scale and the order of emergence found in this study. In Table 35, the points on the ACTFL scale have been “translated” into the concrete grammatical structures that are predicted to appear at that time based on the order of emergence found in this study.

Table 35. Morphosyntactic elements corresponding to ACTFL descriptions

Lev	Description	Morphosyntactic elements
NL	formulas only	
NM	formulas + words and phrases	
NH	syntax strongly influenced by L1	L1 word order (predicted by OG)
	can produce expansions of learned material and stock phrases	thematic verbs
IL	syntax is strongly influenced by L1	L1 word order (predicted by OG)
	can ask a few questions	formulaic questions or questions with SVO order (Predicted by RP, not tested in this study)
	can accomplish uncomplicated communicative tasks.	uses subjects, pronouns, copula, coordination
IM	can produce a variety of questions –	Do-SV(O)-? (<i>Do they want food?</i>) Aux SV(O)-? (<i>Are they going?</i>) Wh-SV(O)-? (<i>What they want?</i>) (Predicted by RP, not tested in this study)
	have difficulty linking ideas	uses coordination, but does not use embedded clauses
	have difficulty manipulating time and aspect	possibly past tense
	have inaccuracies in grammar/syntax	may attempt structures, but not successfully produce them
IH	can exchange basic information	<i>because</i> clauses, negation & do-support
	can narrate and describe in major time frames	modals, especially <i>will</i> past tense emerges for some learners
AL	can narrate and describe in all major time frames	modals: <i>will, could, would</i> past tense used (not always accurately), also with copula forms
	control of aspect lacking at times	aspect may emerge with errors in usage
	“grammatical roughness”	learner may still attempt some structures but not successfully produce them
AM	have good control of aspect	aspect used, but learner may demonstrate inaccurate usage
	have ability to narrate and describe in all major time frames	modals: <i>can, could, will, would</i> , possibly others. some past tense
	lack the ability to consistently provide a structured argument	may fail to use some embedded clauses
	can narrate, describe, explain or tell anecdotes	some relative clauses, expletive subjects may appear (especially those with <i>there</i>)
AH	can explain in detail	variety of complement clauses, relative clauses, sentences with expletive subjects
	can narrate fully and accurately in all time frames	all modals use of aspect

	can construct hypotheses, but patterns of error appear.	some adverb clauses, especially those with <i>if</i> or <i>when</i> expletive subjects, including <i>there</i> and <i>it</i>
	may speak abstractly, but prefer to speak concretely	use of modals besides <i>can</i> and <i>will</i>
	may need to use description or narration in place of argument or hypothesis	may fail to use cancel Aux 2 nd
S	can separate main ideas from supporting information	adverb clauses, relative clauses (all kinds), noun clauses
	can explain opinions	Noun clauses (e.g., <i>I think that...</i>) All modals
	can provide structured argument to support opinions.	Noun clauses Adverb 1 st
	can construct and develop hypotheses to explore alternative possibilities; abstract elaborations	Clauses with <i>if, although, until, etc.</i> Cancel Aux-2 nd clauses (wh- noun clauses) Passive Expletive subjects
		Appropriate use of aspect

It is important to realize that these morphosyntactic elements are intended to serve as a guide, not as an absolute order. That is, it is possible that a learner may produce elements in a somewhat different order than is proposed here. Nevertheless, they can serve to inform an interviewer. For instance, it may be difficult for an interviewer to decide if a speaker can “separate main ideas from supporting information” based solely on a holistic impression. However, it is possible to listen for concrete grammatical structures, specifically, a variety of adverb clauses and relative clauses.

The scale shown in Table 35 is a preliminary proposal only. It requires further testing and refinement before it can be implemented. However, it is an initial attempt to develop a rubric which refers to specific linguistic features and yet is feasible for use as a placement/proficiency test. Questions which remain about its form are elaborated in the next section.

5.2 DIRECTIONS OF FUTURE RESEARCH

The scale proposed in the previous section is preliminary only. Ideally, it should be tested with a large number of learners. The following questions remain:

1. Does the pattern of development outlined in this study apply solely to the kind of data collected in RSA-type activities, or can it also be used to describe other types of oral data (purely spontaneous or elicited)? Additionally, it will be especially important to learn to what extent the pattern applies to various forms of written data or to learners' performance on grammatical discrete-point tasks (e.g., fill-in-the-blank activities, multiple choice tests). The reason that these studies would be particularly important is that many researchers assume that written data (in any form) reflects metalinguistic knowledge and not "true development" (e.g., Young-Scholten, Vainikka & Ijuin, 2005). However, it is not yet known to what extent this is the case, and in any case, it may be useful to measure a student's metalinguistic knowledge as an aspect of proficiency.

2. An additional issue about the modified rubric is whether morphological elements, especially tense and agreement, should be included or kept separate. Research has demonstrated that learners' morphological marking may be dissociated from their syntax, but language testers may (understandably) want to include tense and agreement marking in proficiency measures. It may be possible to score morphological marking separately, as an overall accuracy score.

3. On a related topic, it would be useful to investigate whether accuracy can be incorporated into the rubric. The current ACTFL scale makes many references to learner accuracy. It would be worthwhile to conduct a similar study to this one in which accuracy plays a larger role, to discover whether a developmental table can be produced which refers to accuracy and not just emergence.

4. It is also important to include question development in this study, so that the place of questions in the rubric can be confirmed. Little question data was available for use in this study, despite the fact that both OG and RP make use of them in their predictions. A similar study to this one should be undertaken in which question data is deliberately elicited.

5. What are the causes for the order of development found here? As I mentioned earlier, Pienemann (2003) has been criticized because there is little theoretical backing for aspects of Processability Theory, in particular the Multidimensional Model. The pattern of development proposed here suffers from a similar predicament. However, it would be highly worthwhile to investigate the reasons for the order of emergence found here. The two issues mentioned earlier in the discussion—semantic complexity of form/meaning matching and structural displacement or “gaps”—may be a good place to begin.

6. Can the rubric be used to inform teaching? Pienemann’s Teachability Theory (e.g., 1984, 1985) suggests that learners should be exposed to morphosyntactic elements that are at a stage one above the current stage. It would be highly worthwhile to discover whether this principle operates with the order of elements presented here. It may be the case that the order of emergence presented here also represents the most efficient order of instruction.

7. An important final question involves whether it is possible to connect measures of morphosyntactic development to measures of fluency, pronunciation, and vocabulary use—all of which are included in ACTFL descriptors. That is, morphosyntactic development may occur completely separately from these other areas. Unfortunately, at present it may be impossible to avoid this “mixed measure” approach.

5.3 RECOMMENDATIONS FOR INSTRUCTION AND ASSESSMENT

Some of the limitations of this study are due to the nature of the assessment approaches taken by the English Language Institute at the University of Pittsburgh. That is, it was not possible to investigate the development of question forms by these learners. Similarly, there was often little context for tense and especially 3rd person singular agreement marking. For now, it is recommended that programs incorporate question forms, tense, and 3rd person singular agreement when measuring oral language proficiency. Questions are particularly important, as the variety of English question types require both IP and CP elements, displacement and “gaps,” and various degrees of processing capacity. Unfortunately, at the time, questions are often left out of oral production measures.

It is hoped that the results of this study has begun the process of constructing a proficiency rubric that includes construct-oriented morphosyntactic elements. In doing so, we can work towards bringing together the fields of applied linguistics and language testing, a collaboration that is sure to be a fruitful one.

APPENDIX A

RSA QUESTIONS

2064 Spring 2006

RSA #1 : Topics vary; they are listed below.

Level 2A Talk about a famous person, place or event in your country

Level 3B What sports do you enjoy?

Level 3C Talk about your panel discussion topic: Weddings in your country; Education in your country

Level 4M How do people in your country feel about pets?

Level 4P Describe your country or an interesting place in your country

Level 4Q How do people in your country feel about pets?

Level 5S What experience or opportunity would you like your children to have.

Level 5T Phone your friend Terry to issue an invitation to a party. Leave a message on the answering machine.

RSA # 3

Talk about an important event that happened in the past in your country.

RSA # 4

Talk about a place that you really like. Describe it and tell why you like it.

2067 Summer 2006

RSA #1 : Topics vary; they are listed below.

Level 2A: Talk about a famous person, place, or event in your country.

Level 3B: What sports do you enjoy?

Level 3C: Talk about your panel discussion topic: Weddings in your country or education in your country

Level 4M: How do people in your country feel about pets?

Level 4P: Describe your country or an interesting place in your country.

Level 4Q: How do people in your country feel about pets?

Level 5S: What experience or opportunity would you like your children to have?

Level 5T: Phone your friend Terry to issue an invitation to a party. Leave a message on the answering machine.

RSA #2

Talk about a funny or scary experience that you had.

RSA #3

Talk about your favorite holiday.

2071 Fall 2006

RSA #1

How do you feel about pets? Do many people have pets in your country? How are they treated, in general?

RSA #2

Talk about a person who was very important to you in the past. Who was this person? Why was this person important to you?

RSA #3

What is the biggest problem your country is facing today? How would you change it?

2074 Spring 2007

RSA #1

Is shopping for food in your country the same as in the US? Explain how it is different and how it is the same.

RSA#2

Describe something that you liked to do when you were in your country but that you can't do here. Where did you do this? Why did you like it? How did it make you feel?

RSA#3

Choose a custom (baby's birth, wedding, funeral, entry to adulthood, etc.) in your country. Describe what is done for this custom and why.

APPENDIX A

SAMPLE DATA (L1 = CHINESE)

RSA Level 3: RSA 1 of 4

In my country, China, many people likes badminton in the morning like too, my husband like too. In the Pittsburgh, my husband and I go to the field in the university of Pittsburgh in weekends, we play badminton about two hours, I want to play badminton to kill my health and keep my slim. In my country, I usually play the badminton in the ground of universitys of Beijin, I have the close friend, I have close friend, in the Beijin, and then...and they have...and they like badminton too so we often go...we often play badminton university or college. Stop.

RSA Level 3: RSA 2 of 4

Where do you want to go on vacation? Um..i'm wanted to...I want to some beach, like Hawaii, like, I don't know where. I just want to a beach, because I love swimming in the beach, I like...I like to...I like to see a the ocean, see ocean. The other reason I want to a beach, the other reason is there are some some some fresh fish in there. Maybe I can taste, I can taste some fish in there, I like fish. You know, in Pittsburgh some fish is not fresh, he is f-, they is, they are freezer, but not living.

RSA Level 3: RSA 3 of 4

In my country, I think Beijing applied the Olympic game, that is important event, because in 1992 Beijing applied Olympic game one time, but he is lost, he was lost, Xinee was succeeded. He host-st 2000 olympic game, when Beijing applied 2008 Olympic games, and he was succeeded. I think it is a great time for Beijing, because the China government want to provide China become stronger more and more, and people in Beijing wanted to want to welcome the other country people to visit Beijing.

RSA Level 3: RSA 4 of 4

I think I really like Beijing, because Beijing is captials of China. There have great food and great place, and great spot. I like Chinese foods, for example, dumpling, and fish, and shrimp and anythings. I like Beijing because there has hot spot like Great Wall, and lake, and mountain. I like Beijing because there has my close friend, she help me many things in my career. I like Beijing because I lived here for five years. There has many remain in my hand. The other reason is Beijing will host the Olympic game. I like the food, I like the place, I like the peoples. I think I like Beijing, the reason is there have cheap product products, and cheap food, and fresh vegetable, and fresh shrimp. I like Beijing because I have good job in Beijing.

RSA Level 4: RSA 1 of 3

Today, I want talk my nephew stories. My nephew have a dog. He is sixty years old, six years old. When he had a dog, the dog just six month old. They are good friend. Someday the dog have a pee on the floor. My nephew didn't know. So, he walked the dog and he touch the pee. So his sock is wet, his sock is wet. My nephew blame his dog: "You are bad dog. You are...I didn't...I don't like you". The dog watch my nephew. I think the dog didn't understand my

nephew...what' s my nephew said. My nephew just look at the dog and blame the dog. The dog peered the tail...his tail. I think the dog didn't understand what my nephew said. So he...it is always...

RSA Level 4: RSA 2 of 3

Today I want to talk about my first speech face to thousands students at my college. I remembered, when I...when I...when I came my college, my teacher told me you can prepare a speech face to the thousands student. Ah you will do ah...you will do the speech very well so you can prepare that. That is my first speech face to the thousands students. So I think at that time I am very very nervous. So I write down some words on the paper. That is my first speech. I want...I wanted to tell the student I am ready for the college life, so I...at my speech, I took my paper on the...to the...I took my paper on the table. I read the paper so fast because I am very nervous, when I finish my speech, I go back, I went to back my seat. My classmates ask me what are you...what did you said. I can't understand what did you said because I said so fast. I think many students can't understand me, so that 's my funny experience.

RSA Level 4: RSA 3 of 3

Hi, I think my favorite holiday in China is new year festival, because this is an important holiday in China. Most Chinese people during this holidays always do important things like cook delicious foods and...entire family always gather together. hu. When I was a child, I always get some gift in this famous holiday...in this new years festival. My parents don't ask me to work and don't ask me to study anythings. I always have many fun. In this the new year festival child can play firework, and child can wear the new dress. I think...but now I stay in the U.S., I think my favorite holiday is Christmas, Thanksgiving, because the stores always cut down their price. I can buy some good stuff in the Thanksgiving.

RSA Level 5: RSA 1 of 3

I don't really ... I don't really like pets um...because I scared um....about the pets. When I was a young girl, the...a dog dark forward to...forward...forward me. So...at that time I scared about the pets, but in my country most most people who live in...who live in... who lived in the city have...have a many.... most people who live in the city have many pets like dogs and cats. Um...they treat the pets like her family member...like their family member. They really like them. Um...They always walk around with pets and give the...give the pets a...wonderful food um...as the pets like. um..I have... I have a nephew. He has two dogs in the past. Um...He um...He have two dogs and he really like it. Um...He always play with dogs um...and talk with dogs, um...and he...When he...when his dog lost, he was very sad. Um...I think because they just...my sister just have one child.

RSA Level 5: RSA 2 of 3

The most important person for me is my extra boss. Um. She name...her name is Grace. Um... When I finished my university, I get a bachelor degree...I got my Bachelor degree. Um...I want to... I wanted to got a wonderful job. Um... So I moved to Beijin. Umm...Grace is my first...Grace was my first boss in Beijin. Umm...She teach me....he taught me a lot of skills. Um. That is...like is how to organize a special event. Um. Also...also when she moved to the Shaihai, the...which is the other big city in China, I followed...I followed her to Shaihai because I wanted to learned more experiences from her. He taught...he taught me a lot of how to organize a special event and how to conversation with a people, a lot of business skills. Umm. Also she liked my older sisters, he help me in a...a...strange city how to live, how to find a...how to find

a...how to find a great restaurant...how to find...

RSA Level 5: RSA 3 of 3

The population of China is one of biggest problem now because umm...the biggest problem...the biggest...the biggest populations made us harder to study, harder to go to university or college, hard to get a job. Also make us shortage the use...the sources of shortage. Umm...In our country, um...the regi...the regi...the religion's option is if you have umm...a...many...have several child, that means you are umm...very.umm...very wells. But now something changed in my country, the youngest people...the younger people don't like have too much...have many child in them...in their life. May be one is ok. Also some younger people don't like have any more child. They just enjoy them live. However, in China the government asked us just have one child. I think this is good for China , good for the whole worlds in the future. We can save some space, or save some resource.

APPENDIX B

INTERVIEW QUESTIONS

Introductory Statement:

Thank you very much for agreeing to participate in my study. You will receive \$20 for your time today. I am going to ask you a number of questions. The purpose of these questions is to help me learn about your experience learning English. All of your answers are completely anonymous, which means that I will not use your name in my study. If you don't want to answer one of these questions, that's okay. We can skip it and move on to the other questions. If you change your mind and decide that you don't want to have your answers in the study, you can quit and I will not use your information.

Do you have any questions for me before we start?

Contact questions (modified from Meisel, Clahsen & Pienemann, 1981):

1. How many English-speaking friends do you have?
2. When you're not in class, do you mostly speak Arabic/Chinese? When do you speak English outside of class?

3. Do you participate in ELI activities? Can you give some examples of activities that you have enjoyed?
4. Do you read English books/magazines? Do you read for fun? What do you read?
5. Do you watch English TV/movies? Tell me some of your favorite movies or TV programs.

Attitude toward English-speaking countries

1. What do you think about the United States? What do you think about the role it plays in the world today?
2. Do you think that America and Saudi Arabia/China/Taiwan have a good relationship?
3. Do you like England? Canada? Australia?

Motivation/effort

1. Why do you want to learn English?
2. Do you enjoy learning English? What about it do you enjoy or not enjoy?
3. How much time do you spend doing your homework for your classes? Do you complete all the assignments?

Previous experience

1. Please describe the language classes you took before coming to the ELI. Were they the same or different than your ELI classes?
2. In the past, did you like your English classes? Why or why not?

Self-assessment

1. Are you happy with the progress you have made in the ELI? How have you improved?
2. Which classes do you like the most? (Grammar, speaking, listening, reading, writing)
3. What part of your English do you think you most need to work on?
4. What part of your English do you think you're the best at?
5. Do you think you're ready to go on to academic study in a university? Would additional classes be helpful for you?

Attitudes towards speaking

1. What makes someone a good speaker in your native country?
2. What makes someone a good English speaker?
3. How important is grammar when trying to communicate in English?
4. Do you think you're a good speaker?

REFERENCES

- ACT-ESL website (2007). <http://www.act.org/esl/index.html>. Accessed April 12, 2007.
- Alhawary, M. (2003). Processability Theory: Counter-evidence from Arabic second language acquisition data. *Al-Arabiyya* 36, 107-166.
- ACTFL proficiency guidelines. <http://www.actfl.org/i4a/pages/index.cfm>. Accessed April 5, 2007.
- Alcina, J. & Blecua, J. (1975). *Gramática Española*. Esplugues de Ilobregat: Ariel.
- Alderson, J. (1981). Introduction. In J. Alderson & A. Hughes (Eds.) *Issues in language testing*. London: The British Council.
- Alexopoulou, T., Doron, E. and Heycock, C. (2004). Broad subjects and clitic left dislocation. In *Peripheries: Syntactic edges and their effects*. D. Adger, C. de Cat, and G. Tsoulas (Eds). Oxford: Oxford University Press, 328-358.
- Al-Seghayar, M. (1996). Remarks on the types of complementizers in Arabic. University of Essex manuscript.
- Andersen, R. & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16, 133-156.
- Aziz, Y. (1995). Existential sentences in Arabic-English translation. *Meta* 40, 47-53.
- Bachman, L. & Palmer, A. (1981). The construct validation of the FSI oral interviews. *Language Learning* 31, 67-86.
- Bachman, L. & Palmer, A. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly* 16, 449-65.
- Bachman, L. & Savignon, S. (1986). The evaluation of communicative language proficiency. A critique of the ACTFL Oral Interview. *Modern Language Journal* 20, 380-390.
- Bachman, L. (1988). Problems in examining the validity of the ACTFL Oral Proficiency Interview. *Studies in Second Language Acquisition* 10, 149-164.

- Bachman, L. (1990). *Fundamental considerations in language testing*. New York: Oxford UP.
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.
- Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition* 11, 17-34.
- Becker, M. (2000). The development of the copula in child English: The lightness of *be*. UCLA Doctoral Dissertation.
- Berns, M. S. (1984). Functional approaches to language and language teaching: Another look. In S. Savignon & M. S. Berns (Eds.), *Initiatives in communicative language teaching. A book of readings*. Reading: Addison-Wesley, 3-21).
- Bever, T. (1970). The cognitive basis for linguistic structures. In J. Hayes (Ed.), *Cognition and the Development of Language*. New York: John Wiley and Sons, 279-362.
- Bever, T., & Townsend, D. (1979). Perceptual mechanisms and formal properties of main and subordinate clauses. In W. Cooper & E. Walker (Eds.), *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett*. New York, 159-226.
- Bhatt, R. & Hancin-Bhatt, B. (2002). Structural minimality, CP and the initial state in second language acquisition. *Second Language Research* 18, 348-392.
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning* 33, 1-17.
- Bolinger, D. (1975). *Aspects of language*. New York: Harcourt Brace Janovich.
- Borer, H. & Wexler, K. (1987). The maturation of syntax. In Roeper, T. and E. Williams, (Eds.) *Parameter setting*, 23-172. Dordrecht: Reidel.
- Boser, K., Lust, B., Santelmann, L., & Whitman, J. (1992). The syntax of V2 in early German grammar: The strong continuity hypothesis. *Proceedings of NELS*. Amherst, MA: Graduate Linguistic Student Organization, Department of Linguistics, South College, University of Massachusetts, Amherst.
- Brown, R. & Fraser, C. (1963). The acquisition of syntax. In C. Cofer and B. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1, 1-47.
- Cann, R. (2001). Functional versus lexical: a cognitive dichotomy. In R. Borsley (Ed.) *The Nature and Function of Syntactic Categories*. Burlington: Academic Press, 37-78.

- Carroll, J. (1961). Fundamental considerations in testing for English language proficiency of foreign students. In: *Testing the English proficiency of foreign students*. Washington, D.C.: Center for Applied Linguistics.
- CASAS (Comprehensive Adult Student Assessment System) website (2007). <https://www.casas.org/home/index.cfm>. Accessed April 12, 2007.
- Chang, Charles B. (2005). On existentials, locatives, and the definiteness effect in Korean. Unpublished manuscript, University of California, Berkeley.
- Chapelle, C., Grabe, W. & Berns, M. (1997). Communicative language proficiency: Definition and implications for TOEFL 2000. TOEFL Monograph Series. <http://www.ets.org/Media/Research/pdf/RM-97-03.pdf>. Accessed April 12, 2007.
- Chomsky, N. (1989). Some notes on the economy of derivation and representation. In *MIT Working Papers in Linguistics*. I. Laka & A. Mahajan (Eds). Cambridge: MIT.
- Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.
- Clahsen, H. (1980). Psycholinguistic aspects of L2 acquisition. In S. Felix (Ed.), *Second language development*. Tübingen: Narr, 57-79.
- Clahsen, H., Meisel, J. & Pienemann, M. (1983). Deutsch als Zweitsprache. *Der Spacherwerb ausländischer Arbeiter*. Tübingen: Narr.
- Clahsen, H. (1984). The acquisition of German word order: A test case for cognitive approaches to second language acquisition. In R. Andersen (Ed.), *Second languages*. Rowley: Mass.
- Clahsen, H. (1985). Profiling second language development. In Hyltenstam, K. & Pienemann, M. (Eds.) *Modelling and assessing second language acquisition*. San Diego: College Hill Press, 283-325.
- Clahsen, H. & Muysken, P. (1986). The availability of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research* 2, 93-119.
- Clahsen, H. (1991). Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition* 4, 361-391.
- Clahsen, H., Vainikka, A., & Young-Scholten, M. (1991). Lernbarkeitstheorie und Lexikalisches Lernen. *Linguistische Berichte*, 130, 466-477.
- Clahsen, H., Eisenbeiss, S. & Vainikka, A. (1994). The seeds of structure: A syntactic analysis of the acquisition of case marking. In: T. Hoekstra & B. Schwarz (eds.): *Language Acquisition Studies in Generative Grammar*. Benjamins: Philadelphia, Amsterdam, 85-118.

- Clahsen, Harald, Eisenbeiss, Sonja & Penke, Martina (1996): Underspecification and Lexical Learning in Early Child Grammars. In: Harald Clahsen (ed.): *Generative Approaches to First and Second Language Acquisition*. John Benjamins: Philadelphia, Amsterdam, 129-159.
- Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics* 27, 3-42.
- Clark, J. (1978). Direct testing of speaking proficiency: Theory and practice. Princeton: Educational Testing Service.
- Cohen, D. (1989). *L'aspect verbal*. PUF, Paris.
- Comrie, B. (2007). The acquisition of relative clauses in relation to language typology. *Studies in Second Language Acquisition* 29, 301-310.
- Comrie, B. (1989). *Language universals and linguistic typology*. Chicago: University of Chicago Press.
- Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.
- Crystal, D., Fletcher, P., & Garman, M. (1976). The grammatical analysis of language disability. London: Arnold.
- Davies, A. (1990). *Principles of Language Testing*. Cambridge: Blackwell.
- Deprez, V. & Pierce, A. (1993). Negation and functional projections in early grammar. *Linguistic Inquiry* 24, 25-67.
- Dewaele, J. & Veronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition* 4, 275-297.
- DiBiase, B. & Kawaguchi, S. (2002). Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second Language Research* 18, 274-302.
- Doron, E. and Heycock, C. (1999). Filling and licensing multiple specifiers. In *Specifiers*. D. Adger, S. Pintzuk, B. Plunkett, and G. Tsoulas (Eds). Oxford: Oxford University Press, 69-89.
- Douglas, D. (1988). Testing listening comprehension in the context of the ACTFL proficiency guidelines. *Studies in Second Language Acquisition* 10, 245-261.

- Dube, B. (2000). Where are the Minimal Trees? Evidence from early Zulu L2 subordination. *Second Language Research* 16, 233-265.
- Dyson, B. (1996). The debate on form-focused instruction: A teacher's perspective. *Australian Review of Applied Linguistics* 19, 79-114.
- duPlessis, J., Solin, D., Travis, L., & White, L. (1987). UG or not UG, that is the question: a reply to Clahsen and Muysken. *Second Language Research* 3, 56-75.
- Epstein, S., Flynn, S., Martohardjono, G. (1998). The strong continuity hypothesis: Some evidence concerning functional categories in adult L2 acquisition. In *The Generative study of second language acquisition*. Flynn, S., Martohardjono, G., & O'Neil, W. (Eds), Mahwah: Lawrence Erlbaum Associates, 61-77.
- Ethnologue website (2007). Saudi Arabia. http://www.ethnologue.com/show_language.asp. Accessed April 24, 2007.
- Eubank, L. (1988). Universal grammar, German grammar, and L2 development. Manuscript, University of Texas at Austin.
- Farley, A. & McCollam, K. (2004). Learner readiness and L2 production in Spanish: Processability Theory on trial. *Estudios de Linguística Aplicada* 22, 47-69.
- Fehri, A. (1993). *Issues in the structure of Arabic clauses and words*. Norwell: Kluwer.
- Felix, S. (1992). Language acquisition as a maturational process. In: *Theoretical issues in language acquisition: Continuity and change in development*. Weissenborn, J., Goodluck, H., & Roeper, T. [Eds], Hillsdale, NJ: Lawrence Erlbaum Associates, 25-51.
- Fodor, J., Bever, T., Garrett, M. (1974). *The psychology of language*. New York: McGraw Hill.
- Forster, K. (1979). Levels of processing and the structure of the language processor. In W. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, 27-85.
- Foster, P. & Skehan, P. (1994). The influence of planning on performance in task-based learning. Paper presented at the British Association of Applied Linguistics, Leeds 1994. Quoted in North, B. *The development of a common framework scale of language proficiency*. New York: Peter Lang, p. 53.
- Galal, M. (2005). A minimalist approach to relative clauses in Modern Standard Arabic. Dissertation, University of Kansas.
- Gaya, S. (1961). *Curso superior de sintaxis Española*. Barcelona: Bibliograf.

- Glahn, E., Hakansson, G., Hammarberg, B., Holmen, A., Hvenekilde, A., Lund, K. (2001) Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition* 23, 389-416.
- Goad, H., White, L. & Steele, J. (2003). Missing surface inflection in L2 acquisition: A prosodic account. *Proceedings of the 27th Annual Boston University Conference on language development* 1, 264-275.
- Grimshaw, J. (1990). *Argument structure*. Cambridge: MIT Press.
- Grondin, N. & White, L. (1996). Functional categories in child L2 acquisition of French. *Language Acquisition* 5, 1-34.
- Gruter, T. (2004). Another take on the L2 initial state: Evidence from comprehension in L2 German. *McGill Working Papers in Linguistics/Cahiers linguistiques de McGill* 18, 1-24.
- Guilfoyle, E., Hung, H. and Travis, L. (1992). Spec of IP and spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10, 375-414.
- Habermas, J. (1970). Toward a theory of communicative competence. *Inquiry* 13, 360-375.
- Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. *MIT Working Papers in Linguistics* 21, 275-288.
- Hatch, E. & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Rowley: Newbury House.
- Haznedar, B. (2003). The status of functional categories in child second language acquisition: Evidence from the acquisition of CP. *Second Language Research* 19, 1-41.
- Haznedar, B. & Schwartz, B. (1997). Are there optional infinitives in child L2 acquisition? *Proceedings of the Annual Boston University Conference on Language Development* 21, 257-268.
- Hertel, T. (2003). Lexical and discourse factors in the second language acquisition of Spanish word order. *Second Language Research* 19, p. 273-304.
- Huang, C. (1989). Pro-drop in Chinese: A generalized control theory. In O. Jaeggli & K. Sapir (Eds.), *The Null Subject Parameter*. Dordrecht: Kluwer. 185-214.
- Hudson, T. (1993). Nothing does not equal zero: problems with applying developmental sequence findings to assessment and pedagogy. *Studies in Second Language Acquisition* 15, 461-493.

- Huter, K. (1998). The acquisition of Japanese as a second language. Ph.D. Dissertation. Australian National University.
- Hymes, D. (1971). Competence and performance in linguistic theory. In R. Huxley & E. Ingham (Eds.), *Language acquisition: Models and methods*. London: Academic Press.
- International Second Language Proficiency Ratings (2007). <http://www.islpr.org>. Accessed March 5, 2007.
- Jansen, L. (1991). The development of word order in natural and formal German second language acquisition. *Australian Working Papers in Language Development* 5, 1-42.
- Jarvis, G. (1986). Proficiency testing: A matter of false hopes? *Association of Departments of Foreign Language Bulletin* 18, 22-24.
- Johnston, M. (1985). Syntactic and morphological progressions in learner English. Canberra: Commonwealth Dept. of Immigration and Ethnic Affairs.
- Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21, 60-99.
- Jung, D. (2003). The morphosyntax of complex predicates in Korean and Japanese. Ph.D. dissertation. University of Wisconsin-Madison.
- Kawaguchi, S. (1996). Referential choice by native speakers and learners of Japanese. M.A. Thesis. Australian National University.
- Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge: MIT.
- Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science* 11, 201-258.
- Kim, M. (2007). The absence of the adjective category in Korean. Manuscript, University of Massachusetts. <http://semanticsarchive.net/Archive/TU4NzlkM/MinJoo%20Adjective.pdf>. Accessed April 25, 2007.
- Kim, Y. (1985). The acquisition of Korean. In D. Slobin (Ed.) *The crosslinguistic study of language acquisition*, Vol. 4. Mahwah: Lawrence Erlbaum, 335-437.
- Klein, W., Perdue, C. (1997). The Basic Variety (Or: couldn't natural languages be much simpler?). *Second Language Research* 13, 301-347.
- Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua* 85, 211-258.

- Krashen, S. *Principles and Practice in Second Language Acquisition*. London: Prentice-Hall International, 1987.
- Kuroda S-Y. (1988). Whether we agree or not: A comparative syntax of English and Japanese. In W. Poser (Ed.), *Papers from the Second International Workshop on Japanese Syntax*. Stanford: CSL, 103-143.
- Lakshmanan, U. (1998). Functional categories and related mechanisms. In: *The Generative study of second language acquisition*. Flynn, S., Martohardjono, G., & O'Neil, W. (Eds). Mahwah: Lawrence Erlbaum Associates, 13-16.
- Lantolf, J. & Frawley, W. (1985). Oral proficiency testing: A critical analysis. *Modern Language Journal* 69, 337-345.
- Lantolf, J. & Frawley, W. (1988). Proficiency: Understanding the concept. *Studies in Second Language Acquisition* 10, 181-195.
- Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research* 14, 1-26.
- Lardiere, D. (1998b). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research* 14, 359-375.
- Larsen-Freeman, D. & Long, M. (1991). *An introduction to second language acquisition research*. London & New York: Longman.
- Levelt, W. (1981). The speaker's linearization problem. *Philosophical Transactions*. Royal Society London, 305-315.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Li, C. & Thompson S., (1981). *Mandarin Chinese: A functional reference grammar*. Los Angeles: University of California Press.
- Magnan, S. (1988). Grammar and the ACTFL Oral Proficiency Interview: Discussion and data. *Modern Language Journal* 72, 266-276.
- Michigan Test Website. <http://www.michiganenglishtests.com>. Accessed on April 2, 2007.
- Marinis, T. (2003). *The acquisition of the DP in Modern Greek*. New York: John Benjamins.
- Meisel, J., Clahsen, H. & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition* 3, 109-135.

- Morgan, J., Meier, R., & Newport, E. (1989). Structural packing in the input to language learning: Contributions of intonational and morphological marking of phrases in the acquisition of language. *Cognitive Psychology* 19, 498-550.
- Myles, F. (2004). From data to theory: The over-representation of linguistic knowledge in SLA. *Transactions of the Philological Society* 102, 139-168.
- Newport, E. (1990). Maturation constraints on language learning. *Cognitive Science* 14, 11-28.
- Nisbet, D., Tindall, E., & Arroyo, A. (2005). Language learning strategies and English proficiency of Chinese university students. *Foreign Language Annals* 38, 100-107.
- North, B. (1995). *The development of a common framework scale of language proficiency*. New York: Peter Lang.
- Oller, J. (1976). Evidence for a general language proficiency factor: An expectancy grammar. *Die Neuren Sprachen* 75, 165-174.
- Oller, J. (1983). Evidence for a general language proficiency factor: An expectancy grammar. In: J. Oller (Ed.) *Issues in language testing research*. Rowley: Newbury House.
- Oller, J. (1991). Current research/development in language testing. In S. Anivan (Ed.) *Current developments in language testing*. Singapore: Seameo Regional Language Center, 27-57.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Oxford University Press website. (2007). <http://www.us.oup.com/us/corporate/publishingprograms/esl/titles/tests/qpt/?view=usa>. Accessed on April 12, 2007.
- Penner, Z. & Weissenborn, J. (1996). Strong continuity, parameter setting and the trigger hierarchy: On the acquisition of the DP in Bernese Swiss German and High German. In: *Generative perspectives on language acquisition: Empirical findings, theoretical considerations, and crosslinguistic comparisons*. Clahsen, H. [Ed], Amsterdam, Netherlands: John Benjamins Publishing Co, 161-200.
- Pienemann, M. (1980). The second language acquisition of immigrant children. In S. Felix (Ed.), *Second language development: Trends and issues*. Tübingen: Narr, 41-56.
- Pienemann, M. (1981). *Der Zweitspracherwerb ausländischer Arbeiterkinder*. Bonn: Bouvier.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition* 6, 186-214.

- Pienemann, M. (1985). Learnability and syllabus construction. In Hyltenstam, K. & Pienemann, M. (Eds.) *Modelling and assessing second language acquisition*. San Diego: College Hill Press, 23-76.
- Pienemann, M. (1987). Determining the influence of instruction on L2 speech processing. *Australian Review of Applied Linguistics* 19, 83-113.
- Pienemann, M. & Johnston, J. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.) *Applying second language acquisition research*. Adelaide: NCRC Research Team, 45-142.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics* 1: 52-79.
- Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Philadelphia: John Benjamins.
- Pienemann, M. (2003). Language processing capacity. In Doughty, C. & Long, M. *The Handbook of second language acquisition*. Malden: Blackwell, 679-714.
- Pienemann, M., Johnston, M., & Brindley, G. (1988). constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition* 10, 217-243.
- Pienemann, M. & Mackey, A. (1993). An empirical study of children's ESL development and Rapid Profile. In P. McKay (Ed.) *ESL Development: Language and literacy in schools*. Canberra: Commonwealth of Australia and National Languages and Literacy Institute of Australia, 115-259.
- Pienemann, M. (2003). Rapid Profile stages. <http://groups.uni-paderborn.de/rapidprofile/docs/Stages.pdf>. Accessed May 1, 2007.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge: Harvard University Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge: MIT Press.
- Pollock, J. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365-424.
- Prochazka, T. (1988). *Saudi Arabian dialects*. New York: Kegan Paul International.
- Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Basil Blackwell.

- Rapid Profile. http://groups.uni-paderborn.de/rapidprofile/docs/English_Flyer.pdf. Accessed on April 6, 2007.
- Rizzi, L. (1997). The fine structures of left periphery. In L. Haegeman (ed.) *Elements of Grammar*. Norwell: Kluwer Academic Publishers, 281-337.
- Roeper, T. (1992). From the initial stage to V2: Acquisition principles in action. *The Acquisition of Verb Placement*, Ed. J. Meisel, 333-370. Dordrecht: Kluwer.
- Rule, S. & Marsden, E. (2006). The acquisition of functional categories in early French second language grammars: The use of finite and non-finite verbs in negative contexts. *Second Language Research* 22, 188-218.
- Savignon, S. (2001). Communicative language teaching for the 21st century. In M. Celce-Murcia (Ed.) *Teaching English as a second or foreign language*. Boston: Heinle & Heinle, 13-28.
- Savignon, S. (2002). *Interpreting communicative language teaching: Contexts and concerns in teacher education*. New Haven: Yale University Press.
- Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory* 23, 53-101.
- Schumann, J. (1978). The acculturation model for second language acquisition. In R. Gingras (Ed.) *Second Language Acquisition and Foreign Language Teaching*. VA: Center for Applied Linguistics, 27-50.
- Schütze, C. (2000). Semantically empty lexical heads as last resorts. In: *Semi-lexical Categories: On the content of function words and the function of content words*. N. Corcoran & H. v. Riemsdijk, Eds. Berlin, Mouton de Gruyter.
- Schwartz, B. & Tomaselli, A. (1990). Some implications from an analysis of German word order. In W. Abraham, W. Kosmeijer, & E. Reuland (Eds.) *Issues in Germanic syntax* (251-274). New York: Mouton de Gruyter.
- Schwartz, B. & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12, 40-72.
- Schwartz, B. (1998). On two hypotheses of “transfer” in L2A: Minimal Trees and absolute L1 influence. In: *The Generative study of second language acquisition*. Flynn, S., Martohardjono, G., & O’Neil, W. (Eds). Mahwah: Lawrence Erlbaum Associates, 35-59.
- Schachter, J. & Rutherford, W. (1979). Discourse Function and Language Transfer. *Working Papers in Bilingualism* 19, 1-12.

- Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition*. Rowley: Newbury House, 137-174.
- Smith, C. (1997). *The parameter of aspect*. Dordrecht: Kluwer.
- Snape, N. (in press). L2 acquisition of definiteness and specificity in English by advanced Japanese and Spanish learners. Proceedings of the Generative Approaches to Language Acquisition Conference 2005. Cambridge: Cambridge Scholars Press, 591-6.
- Sohn, S. (1995). *Tense and aspect in Korean*. Honolulu: Center for Korean Studies.
- Song, J. (2005). *The Korean language*. New York: Routledge.
- Spolsky, B. (1989). *Conditions for second language learning: Introduction to a general theory*. New York: Oxford UP.
- Stowell, T. (1978). What was there before there was there? *Proceedings of the 14th regional meeting of the Chicago Linguistics Society*.
- Stoynoff, S. & Chapelle, C. (2005). *ESOL tests and testing*. Alexandria: Teachers of English to Speakers of Other Languages.
- Stromswold, K. (1994). *Lexical and functional categories*. Cambridge: MIT Press.
- Suleiman, S. (1984). Subject-object relations and word order in Standard Arabic. *Proceedings of the Annual Linguistics Conference Irbid, Jordan*.
- Taylor, D. (1988). The meaning and use of the term “competence” in linguistics and applied linguistics. *Applied Linguistics* 9, 148-168.
- TOEFL website. (2007). <http://www.ets.org/portal/site/ets/menuitem>. Retrieved April 2, 2007 from source.
- TOEFL iBT Tour. (2007b) http://www.ets.org/Media/Tests/TOEFL/TPOFlashTour/Container_mx.html. Retrieved April 2, 2007 from source.
- TOEIC website (2007). <http://www.ets.org/portal/site/ets/menuitem>. Retrieved April 2, 2007 from source.
- Tomaselli, A., & Schwartz, B. (1990). Analysing the acquisition stages of negation in L2 German: Support for UG in adult SLA. *Second Language Research* 6, 1-38.
- Towell, R. & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon: Multilingual Matters.

- Upshur, J. (1971). Objective evaluation of oral proficiency in the ESOL classroom. *TESOL Quarterly* 5, 47-59.
- Vainikka, A. & Young-Scholten, M. (1994). Direct access to X'-Theory: Evidence from Korean and Turkish adults learning German. In *Language acquisition studies in generative grammar*. Hoekstra, T., & Schwartz, B. (Eds), Amsterdam, The Netherlands: John Benjamins Publishing Company, 265-316.
- Vainikka, A., & Young-Scholten, M. (1996a). Gradual development of L2 phrase structure. *Second Language Research* 12, 7-39.
- Vainikka, A., & Young-Scholten, M. (1996b). The early stages in adult L2 syntax: Additional evidence from Romance speakers. *Second Language Research* 12, 140-176.
- Vainikka, A., & Young-Scholten, M. (1998a). The initial state in the L2 acquisition of phrase structure. In: *The Generative study of second language acquisition*. Flynn, S., Martohardjono, G., & O'Neil, W. (Eds). Mahwah: Lawrence Erlbaum Associates, 17-34.
- Vainikka, A., & Young-Scholten, M. (1998b). Morphosyntactic triggers in adult SLA. In: *Morphology and its interfaces in second language knowledge*. Beck, M. (Ed.) Amsterdam: John Benjamins, 89-113.
- Vainikka, A., & Young-Scholten, M. (2002). Restructuring the CP in L2 German. *Proceedings of the Annual Boston University Conference on Language Development* 26, 712-722.
- Vainikka, A. & Young-Scholten, M. (2006). The roots of syntax and how they grow: Organic Grammar, the Basic Variety, and Processibility Theory. In S. Unsworth, T. Parodi, A. Sorace & M. Young-Scholten (Eds.) *Paths of development in L1 and L2 acquisition*. Amsterdam: Benjamins, 77-106.
- Vollmer, H. (1981). Why are we interested in general language proficiency? In J. Alderson & A. Hughes (Eds.) *Issues in Language Testing*. London: The British Council, 152-171.
- Vollmer, H., & Sang, F. (1983). Competing hypotheses about second language ability: A plea for caution. In J. Oller (Ed.) *Issues in language testing research*. Rowley: Newbury House, 29-79.
- Weissenborn, J. (1990). Functional categories and verb movement: The acquisition of German syntax reconsidered. *Linguistische Berichte* 3, 190-224.
- White, L. (1989). *Universal Grammar and second language acquisition*. Amsterdam: John Benjamins.
- White, L. (1991). Second language competence versus second language performance: UG or processing strategies? In L. Eubank (Ed.), *Point-counterpoint. Universal Grammar in the second language*. Philadelphia: John Benjamins, 67-189.

- White, L. (2002). Morphological variability in endstate L2 grammars: The question of L1 influence. *Proceedings of the annual Boston University conference on language development* 26, 758-768.
- White, L. (2003a). *Second language acquisition and universal grammar*. New York: Cambridge University Press.
- White, L. (2003b). Fossilization in steady-state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6, 129-141.
- White, L., Valenzuela, E., Kozłowska-MacGregor, M., & Leung, Y. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25, 105-133.
- Woolford, E. (1991). VP-internal subjects in VSO and nonconfigurational languages. *Linguistic Inquiry* 22, 503-540.
- Wymann, A. (1996). *The expression of modality in Koran*. Ph.D. Dissertation. University of Bern.
- Young, R. (1995). Conversational styles in language proficiency interviews. *Language Learning* 45, 3-42.
- Young-Scholten, M., Ijuin, C., & Vainikka, A. (2005). Organic grammar as a measure of development. TESOL. San Antonio.
- Young-Scholten, M. & Ijuin, C. (2006). Expanding Organic Grammar-based assessment. TESOL. Tampa.