2007-07-13

A Web-based Tool for Oral Practice and Assessment of Grammatical Structures

Heather Colleen Torrie
Brigham Young University - Provo

A WEB-BASED TOOL FOR ORAL PRACTICE AND ASSESSMENT OF GRAMMATICAL STRUCTURES

by

Heather Torrie

A selected project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Linguistics and English Language

Brigham Young University

August 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a selected project submitted by

Heather Torrie

This selected project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

_____	_____
Date	Diane Strong-Krause, Chair
_____	_____
Date	Mark W. Tanner
_____	_____
Date	C. Ray Graham

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the selected project of Heather Torrie in its final form and have found (1) its format, citations and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures. tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

_____ _____

Date Diane Strong-Krause
Chair, Graduate Committee

Accepted for the Department

_____ _____

Date William Eggington
Department Chair, Department of Linguistics and English Language

Accepted for the College

_____ _____

Date Gregory D. Clark
Associate Dean, College of Humanities

ABSTRACT

A WEB-BASED TOOL FOR ORAL PRACTICE AND ASSESSMENT OF GRAMMATICAL STRUCTURES

Heather Torrie

Department of Linguistics and English Language

Master of Arts

The grammar course objectives at Brigham Young University's English Language Center (ELC) are a list of grammar structures for each level that students are expected to be able to master. These objectives currently are only measured by pencil-paper grammar tests, providing information regarding receptive grammar ability only. Therefore, there is a need for an oral grammar assessment to measure productive grammar ability, providing diagnostic and achievement information about the specific grammar objectives.

This project is a web-based oral grammar assessment tool that enables teachers to assess students' mastery of the structures covered in the ELC grammar classes. The core of the project is an online database of speaking tasks designed to target specific grammar structures. Teachers can access the database through a website and create

assessments using groups of selected speaking tasks. Then, students access the speaking tasks through a separate recording application that has been developed to display the prompts and record students' responses. Teachers can access the recorded responses on the website and rate the responses, using a rubric to measure students' mastery of each grammar structure.

An evaluation of the project reveals that students and teachers feel that it is beneficial in providing practice and self-assessment opportunities. Most students like using the program and consider it helpful. However, a major challenge is the time required of teachers to listen to students' responses in order to use the program for actual assessment and give effective feedback. As the program is used more widely, future research needs to investigate the reliability and validity of the tasks and using the rubric for assessment.

ACKNOWLEDGEMENTS

I would like to express my appreciation to my committee for their patience, as well as to the teachers and staff at the English Language Center for their assistance as I developed this project. I also wish to express my gratitude to my brother, Michael, for help and support through the entire process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

BACKGROUND

During the Fall of 2005, I taught a high-intermediate grammar class at the English Language Center (ELC). This was my first attempt at teaching grammar and it was somewhat overwhelming. As we worked through some of the more complicated structures, such as the past perfect tense and the unreal conditionals, I couldn't help but notice a discrepancy between the students' written exercises and their everyday speech. For instance, students could produce the complex grammar forms we studied in a fill-in-the-blank homework assignment, but they made frequent errors when speaking, or avoided the structure altogether.

I also noticed the discrepancy between students' performance in written grammar exercises and their performance on the ELC speaking tests. At the ELC, each skill area has a Level Achievement Test (LAT) at the end of the semester. I had the opportunity to rate the speaking LAT several times, and similar to my observation of students' everyday speech, I found that their LAT speaking responses were often lacking in appropriate and high-level grammar. It seemed that high-intermediate students could complete the assigned speaking task by using simple grammar structures taught in the beginning level grammar course. For example, one of the high-intermediate speaking objectives is to give and support an opinion. Many students could accomplish this task by using the simple modals of advice such as *should* and *must*, which are taught in the high-beginning grammar class.

Students' struggle to use accurate and appropriate-level grammar in speaking, and also their avoidance of particular structures, demonstrates the need for oral grammar

assessment. Currently, grammar assessment at the ELC has been strictly the pencil-paper quizzes and tests. This kind of testing only measures receptive grammar ability, except for some limited production involved in sentence-writing items. Therefore, there is a need to assess productive grammar ability in speaking.

To assess students' productive grammar ability, I originally explored the possibility of giving a separate grammar score on the speaking LAT. However, I found that a major concern was that raters need to consider many other factors besides grammar, such as fluency, pronunciation, phrases and vocabulary, and content. Furthermore, raters are not always trained in grammar and therefore it is difficult to give a grammar production score, as well as feedback that is specific to the structures being learned in the grammar classes. Because grammar is not easily assessed separately on the speaking LAT, there is a need for a separate oral grammar assessment that could provide information to ELC grammar teachers and students regarding the course objectives.

Given the opportunity to fill this need, I decided to develop an oral grammar assessment tool to focus specifically on students' mastery of the structures studied in their grammar classes. The assessment tool consists of an online database of speaking tasks targeting specific structures, together with an application to record the responses. The evaluation of students' recorded responses will help teachers and students know which structures have been mastered in speaking and which need more attention.

CHAPTER 2

LITERATURE REVIEW

When contemplating grammar testing, several questions arise: What is grammar knowledge, and what is the rationale behind teaching it? How should grammar be taught in order to develop implicit knowledge? What do researchers suggest to effectively test grammar? How has research addressed classroom assessment? Because my project involves developing and implementing a classroom grammar assessment tool, this literature review aims to answer these questions. The first two sections define the construct of grammar and discuss the methods of teaching grammar suggested by research. The final section addresses grammar assessment by first giving an overview of general classroom assessment, and then discussing frameworks of test development specifically for grammar.

Defining the Construct of Grammar

One of the first steps in developing assessment instruments is to define the construct, which strengthens the construct validity in an assessment. With a well-defined construct, a test can be designed and used to make appropriate inferences (see Messick, 1989). In defining the construct of grammar, it is necessary to examine how research has categorized grammar and the role it plays in language acquisition. Most bodies of grammar research deal in some way with the dichotomy of implicit versus explicit learning, and the role of implicit and explicit knowledge.

Implicit and Explicit Knowledge

Implicit grammar knowledge is commonly defined as the ability to use a structure intuitively (Purpura, 2005). It also refers to grammatical performance, or the actual use of implicit knowledge in language production activities (Hymes, 1972). In SLA research, a defining feature of implicit knowledge is automaticity. In other words, knowledge that is truly acquired is available for automatic use (DeKeyser, 2003).

Conversely, explicit grammar is often referred to as grammatical knowledge, or knowledge about the rules and structures of a language (Purpura, 2004). Ellis (2004) defines explicit grammar as "knowledge *about* language and *about* the uses to which language can be put" (p. 229). He proposes that it is a conscious awareness of some feature of the language. Ellis (2006) also explains that there are two types of explicit knowledge: analyzed knowledge is the awareness of how a grammar structure works, while metalinguistic knowledge, includes the metalanguage, which helps to verbalize grammar explanation and rules.

The idea that explicit knowledge is entirely different from implicit knowledge is sometimes known as the non-interface position (Ellis, 2006). For instance, Krashen (1982) argues that the gap between explicit knowledge and use is too great to ever be beneficial. However, some researchers believe that explicit knowledge is of some value in facilitating the development of implicit knowledge, which is referred to as the interface or weak interface position (Ellis, 1994, 2006). DeKeyser (1997, 1998, 2003) believes that if learners have enough practice, explicit knowledge can become proceduralized and be as accessible as implicit knowledge. He also points out that there is no evidence in research that explicitly learned knowledge cannot become automatized.

Implicit and Explicit Learning

It is also important to understand the different processes by which implicit and explicit knowledge are learned. DeKeyser (2003) states that the defining feature of implicit learning is the lack of awareness. The best example of implicit learning is that of children acquiring their first language, in which the learners are completely unaware of what they are learning, and the language becomes automatic. Explicit learning, on the other hand, is learning with an awareness of the linguistic features being learned. Usually, this involves an explanation of the rules and structures. DeKeyser explains that knowledge gained through explicit learning could possibly become implicit if the awareness of the structure is lost over time. Likewise, implicitly learned knowledge could become explicit if the learner ever becomes aware of the particular structure.

Many studies have attempted to compare the effects of these different types of learning. Most laboratory studies that directly compare the effect of explicit instruction with implicit instruction show a benefit to explicit instruction (Doughty, 1991; Robinson, 1996). Norris and Ortega's (2000) meta-analysis shows that classroom studies also suggest a benefit to explicit instruction. However, DeKeyser (2003) argues that many of the studies are not true comparisons of implicit/explicit learning because they did not involve true implicit learning. He explains that in those studies it was impossible to ensure that the participants were completely unaware of the grammar rules. For example, it could be argued that the participants in VanPatten and Oikkenon's (1996) study were actually learning explicitly through an inductive method, and were being compared with an explicit deductive group. Despite these limitations, research suggests at least some benefit to explicit learning.

Summary of Grammar Definitions

The construct of grammar is a dichotomy of implicit and explicit knowledge acquired by implicit or explicit means, and most researchers agree that language acquisition means fully automatic, implicit knowledge. Therefore, grammar ability is the ability to automatically use correct grammar structures. Some researchers believe that implicit knowledge can be acquired through explicit learning, just as implicitly learned knowledge can become explicit if brought to attention. Lastly, some research has compared explicit and implicit learning, and a group of studies suggests a benefit to explicit learning.

Teaching Grammar

The role of instruction in SLA has long been debated (Krashen, 1982; Long, 1988). The major objection to instruction is the theory of acquisition order, which is the belief that grammar is acquired in a natural order, regardless of instruction (Larsen-Freeman & Long, 1991, sec. 4.3). However, the majority of research suggests there is a benefit to instruction, one example being an increase in the rate of acquisition (DeKeyser, 1998). This is primarily because of the Noticing Hypothesis, which is the idea that learners must notice specific aspects of the language in order to acquire them (Doughty, 2003). In some cases, learners are not able to notice less salient, complex structures without assistance. The purpose, then, of grammar instruction is to help learners notice by drawing their attention to particular structures and forms.

Focus on Form

The purpose of focus on form instruction is to help students notice and correct the gap between what they hear and see and what they produce. In their study, Williams and Evans (1998) found that while a focus on form is beneficial, its effectiveness depends on the type of form. Explicit instruction about forms that are task-essential, such as participle-adjectives, was more effective than flooded input alone using that structure. On the other hand, structures that are less essential to communicating meaning, such as the passive voice, were not as effectively learned by focus on form. The complexity variable also plays a role in determining which structures are most effectively taught. According to DeKeyser (1998), complex and abstract structures cannot be acquired without a strong focus on form and error correction, whereas simple and non-abstract rules can be learned implicitly. Determining the complexity of a form, however, is sometimes difficult because researchers sometimes disagree on whether a rule is simple or complex.

Learner readiness also plays a role in the effectiveness of focus on form. This means that learners must be developmentally ready to acquire the form, especially when the form is complex. When learners have already noticed a structure and are beginning to use it, the explicit instruction and continuous error correction helps to solidify the form. On the other hand, if they have not noticed the structure before the instruction, the focus on form instruction is less effective (Williams & Evans, 1998).

Practice

If a particular grammar form is to be useful, it must become automatized. As Anderson & Fincham (1994) explain it, declarative knowledge, the ability to explain the

grammar rules learned through focus on form instruction, can then become procedural knowledge when a behavior is also involved. To reach full proceduralized knowledge, which is automatic, learners must engage in communicative and meaningful practice, allowing learners to access their declarative knowledge from their working memory. This kind of practice allows the link between form and meaning to solidify and become part of their long-term memory (DeKeyser, 1998).

Task-Based Teaching

Communicative language teaching, which has been the trend in the second language classroom over the past few decades, is aimed at providing communicative practice which could then facilitate the proceduralization of explicit knowledge. More recently, there has been an emphasis on task-based language teaching (Kumaravadivelu, 2006). A task can be defined as a workplan with input and instructions requiring learners to attend to meaning, while at the same time using authentic communication, pragmatics, and interaction (Ellis, 2003). Furthermore, tasks often involve real-world activities, with a focus on the language outcome (Skehan, 1998).

SLA research suggests that a major benefit of task-based language teaching is the opportunity for learners to negotiate meaning (Ellis, 2003). Long's (1983) Interaction Hypothesis emphasizes that the best comprehensible input is provided when learners are required to negotiate for meaning. This comprehensible input can then lead to acquisition through the negotiation provided in task-based language teaching, where learners need to interact to complete the task together.

Ellis (2003) distinguishes between two types of tasks. Unfocused tasks are those which are intended to elicit any natural language that occurs in completing the task, while

focused tasks, on the other hand, are designed to elicit a specific linguistic feature. Ellis notes that there is a distinct difference between a focused task and a situational grammar exercise. In communicative language teaching, grammar exercises are often used to teach particular structures by presenting the rules, followed by focused practice and a communicative exercise. However, task-based teaching allows learners to select which linguistic forms to use. They are not told which grammatical structures they need to practice, although the task can be designed to naturally target a specific feature.

Summary of Grammar Teaching Literature

Most of the research reviewed indicates that there is a benefit to grammar instruction. Focus on form instruction is useful to help learners notice the gap between the language input they receive and their output. Then, with communicative and meaningful practice and interaction provided through task-based instruction, explicit grammar knowledge can become automatized.

Grammar Assessment

Now that the definition of the construct of grammar and the rationale for grammar instruction have been established, it is necessary to review the literature relevant to testing grammar. This section first explores the role of classroom assessment in general, including the purposes and methods for implementing different types of assessment. Next, the focus turns specifically to grammar testing, discussing test development frameworks, performance testing, and scoring rubrics.

Formative and Summative Assessment

It is important to note the difference between summative and formative assessment. Summative tests are used by administration to certify competence or promote students. Formative tests, on the other hand, are generally thought of as the means of providing feedback in order to improve learning and teaching. With new standards in many areas of education, teachers feel pressure to show that students measure up to certain benchmarks (Arkoudis & O'Loughlin, 2004; Brindley, 1998). For this reason, formative assessment often gives way to summative assessment (Rogers, 1991). In second language learning, formal external tests, such as the iBT TOEFL, seem to have a strong influence on both classroom instruction and testing procedures. Studies which surveyed many different kinds of language institutions found that classroom assessment methods often mirrored the relevant external and institutional achievement tests in format and content (Cheng, 1999; Li, 1990).

Despite the strong influence of external tests, there is often a genuine desire to align classroom assessment with desired learning outcomes and use formative assessment to improve instruction and provide useful feedback to the student. The effort to make assessment, particularly formative assessment, consistent with classroom learning and instruction goals is sometimes called learning-oriented assessment (Purpura, 2004). Another issue to consider here is the potential for positive washback, the influence a test can have on learning and teaching. Alderson and Wall (1993) hypothesize that a test can affect what is taught, the rate and sequence of learning, and attitudes toward the content and methodology. By incorporating learning goals and authenticity into classroom

assessment, as well as providing detailed diagnostic information, positive backwash is likely to occur (Bailey, 1996; Shohamy, 1992).

Classroom Assessment Methods

Classroom assessment can include both formative and summative assessment, although most of what happens in the classroom is formative. Cheng, Rogers, and Hu (2004) conducted a study which explored the methods of classroom assessment in various institutions in Canada, Hong Kong, and Beijing. The results of the study showed that most instructors spent almost one quarter of total instruction time on assessment and evaluation. Instructors reported that learning was the main purpose for their assessments. As mentioned above, learning-oriented assessment in the classroom is aimed at providing diagnostic and achievement feedback to students. This kind of assessment also provides motivation to progress, as well as preparation for future high-stake tests. Another purpose for learning-oriented assessment in the classroom is to inform instruction, helping instructors plan what to teach and diagnose their own teaching abilities (Cheng et al., 2004).

The majority of instructors in the study discussed above also reported developing their own assessments and working together with other instructors. A smaller proportion used published test items from sources such as textbooks. The format used in assessments, such as selected-response, limited-, and extended-production tasks, varied according to geographic setting and course context. Concerning feedback and reporting, the most common form was oral feedback and written comments, along with scores.

Grammar Test Development

Applying the principles discussed above to grammar testing specifically, positive washback is likely to occur when grammar tests are aligned with communicative goals and instruction through learning-oriented assessment. This section will now discuss the methods suggested by research of test development, performance testing, and scoring rubrics, serving as a guide to implementing classroom grammar assessment.

The general purpose of a grammar test is to collect information about how well learners use grammar to convey meaning in certain situations (Purpura, 2004). It is important, then, to determine the situations in which learners would use the target language to communicate. According to Bachman and Palmer's (1996) framework of test development, one of the first steps is to identify target language use (TLU) tasks that learners will need to perform in those situations. Therefore, the main goal of test development is to create test tasks that are aligned with TLU tasks, which leads to test authenticity. As TLU tasks are defined, it is also necessary to identify the grammar structures needed to complete the task, providing a theoretical definition of the construct (Purpura, 2004).

With a theoretical definition of grammatical ability established, it is possible to identify what evidence is needed to support that particular claim of ability. Test tasks are then designed to elicit the appropriate kind of evidence. The Bachman and Palmer (1996) framework identifies important characteristics to consider when developing test tasks. First, it is important to consider the setting in which the task is to be performed, including the physical setting, such as a classroom or computer lab, as well as the time period in which the test task will be administered. Other important characteristics are

instructions and time allotment. Ellis (2001) claims that speeded tasks are more effective for measuring implicit knowledge, because examinees do not have sufficient time to access their explicit knowledge. Bachman and Palmer also identified the format and language characteristics as important parts of grammar test tasks. For example, tasks must have some kind of stem or prompt, and directions explaining what is required. Output characteristics of grammar test tasks include the format of the response and the expected language outcome, all of which depend on the type of task.

There are three main types of test tasks seen on traditional grammar tests (Purpura, 2004). Selected response tasks simply require the examinee to choose the correct answer, like multiple choice or matching items. Another type is the limited-production tasks, such as the fill-in-the-blank and short-answer items. Finally, extended-production tasks involve a prompt of some kind, such as information gap exercises, story-telling, role-playing, and simulation tasks.

Performance Testing

A growing group of language testing literature focuses on performance, or task-based, testing. This involves assessing actual performance to complete a task, rather than assessing only abstract knowledge (McNamara, 1996). As Messick (1994) describes it, performance is the vehicle for assessment and reveals something about the underlying language knowledge. The push for performance testing came about mainly to align testing with communicative teaching methodology (McNamara, 1996).

Performance testing often involves giving a prompt and allowing some planning time, the amount of which distinguishes planned from unplanned discourse. Planned discourse allows the learner sufficient time to access explicit grammar knowledge and

produce more complex structures than would be possible without planning time (Ochs, 1979). Conversely, unplanned discourse is spontaneous speech produced strictly from the learner's implicit knowledge. One group of studies looking at the effect of pre-planning on actual performance suggest that longer planning time often correlates with higher fluency and more complex language (Ellis, 2003; Ortega, 1999; Yuan & Ellis, 2003). However, the correlation between planning time and grammatical accuracy is not so apparent in those studies. Ellis (2001) argues that with little or no planning time learners are prevented from accessing their explicit knowledge, thus giving a more accurate measure of their implicit grammatical knowledge. Because of this, it is sometimes believed that extended-production tasks could measure implicit grammatical knowledge more than limited-production tasks (Purpura, 2004).

A challenge that arises in developing performance tasks designed to elicit specific language is the problem of preserving authenticity. Cronbach (1984) calls this the bandwidth fidelity dilemma, meaning that authentic speech samples can give more generalizable information about learners' overall proficiency, but less accurate conclusions about a specific language ability. For example, learners may be able to complete a given speaking task without using a targeted grammar structure. On the other hand, less authentic speech samples are less generalizable, but give more accurate information about a specific language ability. This is a relevant concern when designing test tasks that are authentic but also focused enough to elicit specific grammar structures.

Developing and Using Scoring Rubrics

Responses from a performance or task-based assessment must be measured with a rubric, which can be thought of as the theoretical embodiment of the construct

(McNamara, 1996). In grammar testing, a holistic scale is used to score overall grammatical ability. This type of scale usually encompasses both grammatical accuracy and complexity of grammar. Using a holistic score is sometimes more practical; however, holistic scores are sometimes lacking in diagnostic information telling learners which aspects of grammar they need to improve (Purpura, 2004). Furthermore, because raters often focus on different aspects of the performance, holistic scales are sometimes subject to interpretation problems (Weigle, 2002). Analytic scales are used to focus on particular features of language ability, giving separate scores for grammatical accuracy, meaningfulness, and pragmatic appropriateness (Purpura, 2004).

Weigle (2002) offers guidelines for developing scoring rubrics. The most important consideration is to identify which aspects of language ability are the most important in the assessment, or, what construct the test is trying to measure. Another consideration is the level descriptions. Rubrics need to be usable and interpretable by all stakeholders, including students. In determining the number of points for each level, Bachman and Palmer (1996) recommend using a zero score to designate no evidence of mastery and having the highest score represent full mastery. They suggest a descriptive approach for creating a rubric by the administering the test first, and from the responses, it is possible to identify examples of different levels of mastery. Characteristics of those responses can then be used to write the level descriptions.

Using scales to rate performance always involves subjective judgments. In order to minimize variability among raters, it is important to have clearly worded level descriptions. Rater training on how to use the rubric is also necessary, and research suggests that this helps raters become more self-consistent and reduces extreme ratings,

such as excessive severity or leniency (Shohamy, Gordon, & Kraemer, 1992). It is also a good practice to use double-ratings, with a third rater for dealing with disagreement (McNamara, 1996).

Summary of Grammar Testing Literature

Literature investigating classroom assessment distinguishes between formative and summative assessment, and it is suggested that language instructors spend a lot of energy developing their own formative tests. Furthermore, despite negative washback from external tests, learning-oriented formative assessments can be implemented to provide helpful diagnostic information to language learners and teachers. When classroom assessments are aligned with class goals, tasks are authentic, and feedback is detailed, there is a potential for positive washback. For testing grammar, focused-tasks used in performance testing can be used to elicit and assess implicit knowledge of specific grammatical features. Then, using clearly-defined rubrics that reflect the construct, grammar assessment has the potential for positive washback and can facilitate learning.

Bringing It Together

The literature reviewed in this section brings together the main issues surrounding grammar testing. A close examination of the construct of grammar reveals the dichotomy between explicit and implicit grammar. Because implicit grammar knowledge is required for automatic use in language and communicative activities, grammar instruction and testing must be aimed at eliciting implicit grammar knowledge. It is also important to note that the main benefit of grammar instruction is to help learners notice and correct

their grammar mistakes. Furthermore, classroom activities that teach explicit grammar must also include practice to enable learners to proceduralize, or automatize, their explicit knowledge. As for assessing grammar, frameworks in grammar testing theory involve developing test tasks based on actual language use. Performance testing, particularly task-based testing, uses real-time performance as an attempt to measure implicit knowledge and language ability. Then, aligned with course goals emphasizing meaningful practice, learning-oriented assessment promotes positive washback in motivating students to notice and correct their output in language production.

CHAPTER 3

PROJECT RATIONALE

The literature reviewed in the previous chapter reveals the purpose of grammar instruction and how to assess grammar in the classroom. This chapter will now provide a rationale for my project by demonstrating the need for a separate oral assessment to measure grammar. The first two sections discuss how grammar is currently taught and tested at the ELC, including a description of the course objectives and grammar teaching practices. The final section illustrates how an oral grammar test is supported in research and could benefit learning and assessment at the ELC.

Grammar Teaching Objectives and Methodology at the ELC

The ELC currently has five levels, with Level 1 considered a low-beginning class, Level 2 high-beginning, Level 3 low-intermediate, and Level 4 high-intermediate. Levels 1-4 have four individual skill area classes, including reading, writing, listening/speaking, and grammar; Level 5, however, is strictly content-based with the four skills incorporated into four classes. All classes have a set of established objectives set by the skill area coordinators. For Levels 1-4 grammar classes, the objectives consist of a list of 20-40 grammar structures that are based on those covered in the *Focus on Grammar* series (Fuchs & Bonner, 2006; Fuchs, Bonner & Westheimer, 2006; Schoenberg, 2006; Schoenberg & Maurer, 2006). Written instructions on course objectives state that "students should be able to use the…structures with 80% accuracy" at the end of the semester (English Language Center, 2006a). In theory, this means that students should be able to use the structures in speaking and writing production. However, this objective is

currently only measured through pencil-paper grammar tests (F. Keeler, Personal Communication, February 26, 2007). Teachers are also expected to state their specific class objectives in their syllabi, ensuring that they are consistent with the established ELC class objectives (English Language Center, 2006a). Class objectives often state that students should be able to use or produce the structures in activities such as conversation, writing, and on tests.

Grammar is taught explicitly, usually by introducing the structure and rules deductively, on the chalkboard or with a PowerPoint presentation. Sometimes inductive methods are used, such as having students look at examples and determine the grammar rules together in small groups. Generally, ELC teachers spend 15-25 minutes during each class on the grammar presentation. Then teachers usually provide practice activities, such as workbook exercises and worksheets, followed by communicative activities, giving students the opportunity to produce the structure. Examples of communicative activities used in ELC grammar classes include information gap, description, and role-plays (F. Keeler, Personal Communication, February 26, 2007).

Over the past few years, there has been an emphasis on integrating grammar instruction with the tasks covered in the listening/speaking classes (English Language Center, 2006a). The rationale is that students receive multiple exposure to the listening/speaking vocabulary, and at the same time, integration allows listening/speaking teachers to focus on vocabulary, content, fluency, and pronunciation, rather than to review grammar. In grammar classes, the integration of listening/speaking topics provides contextualized practice, making grammar less abstract with meaningful practice activities focusing on grammar production.

Therefore, it seems that ELC grammar classes are helping students to become aware of the grammar and notice the gap between the input they receive and their own output (see Doughty, 2003). With the meaningful and communicative practice activities used on a daily basis in ELC grammar classes, explicit knowledge can become proceduralized, allowing students to actually produce the grammatical structures they learn (see DeKeyser, 2003).

Assessing Grammar at the ELC

At the ELC, grammar is assessed both formatively and summatively. Teachers are expected to regularly assess students' grammar ability through tests and quizzes (English Language Center, 2006a). Although it is not required, most grammar teachers at the ELC work together to develop and refine classroom tests, and then the same basic test can be used for all classes at the same level. Currently, classes in each level give approximately 5-10 tests throughout the semester to provide formative assessment (F. Keeler, Personal Communication, February 26, 2007). Despite the use of communicative and productive practice activities, the general format for classroom tests is strictly the written kind with multiple choice, error correction, sentence-completion, and short answer. Some tests may include a few short paragraph- or sentence-writing items.

Level Achievement Tests (LAT's) for each skill area provide summative assessment at the end of each semester. The Grammar LAT for each level consists of 75 multiple-choice items of three kinds: fill-in-the-blank, fill-in-the-blank cloze, and error detection. Of these 75 items, the LAT contains approximately two items focused on each grammar structure covered in the course (English Language Center, 2006a). The

grammar proficiency grade is calculated by basing 50% of the grade on classroom assessments and 50% on the LAT.

In the listening/speaking classes, oral grammar production is assessed indirectly as part of speaking proficiency, which is assessed through classroom speaking tests throughout the semester and on the Speaking LAT. Students complete speaking tasks on the computer and are rated according to a feedback sheet that helps the rater give a final score, as well as give useful feedback to the students. The feedback sheet separately addresses the areas of pragmatics, content, detail, discourse level, grammar, pronunciation, and vocabulary and phrases (English Language Center, 2006b). Although this feedback sheet allows teachers to give actual scores for these features, it is only to provide general feedback to the students. After considering each area, the teacher then, almost intuitively, gives a holistic score for speaking proficiency. Figure 1 is an excerpt from the feedback sheet and shows how grammar is addressed in the listening/speaking classes, along with the other areas of speaking proficiency. A rating of 1 means "Needs a lot of improvement" while a score of 4 means "Excellent."

Grammar:	**1**	**2**	**3**	**4**
-	Errors were minimal and not distracting.			
-	Used level-appropriate grammar			
-	Used a variety of grammar structures.			

Figure 1. The section addressing grammar on the speaking LAT feedback sheet.

Thus, in classroom speaking tests and on the LAT, grammar is considered, but only in a general way. Furthermore, the scores given using the grammar rubric on the feedback sheet do not reflect the objectives from the grammar courses.

Need for an Oral Grammar Assessment

Classroom speaking tests and the speaking LAT are a form of performance testing measuring overall speaking proficiency. As mentioned above, however, grammar is only one aspect of oral proficiency incorporated in the speaking rubric at the ELC, and the feedback given is not connected to the objectives of the grammar classes. Because the speaking tasks are designed to measure students' mastery of the speaking objectives, the tasks themselves do not necessarily elicit those particular structures included in the grammar class objectives. Students may avoid certain structures that are difficult, because they are able to adequately complete the task without using them.

Given that the ELC speaking tests cannot specifically elicit particular grammar structures or provide specific diagnostic information relevant to the grammar objectives, there is a need for a separate assessment to measure the mastery of specific structures. A separate grammar assessment would give grammar teachers and students diagnostic and achievement information relevant to those structures covered in the grammar courses.

Another benefit to an oral grammar assessment is the potential for positive washback. According to the literature reviewed, positive washback is likely to occur when assessment is aligned with class goals and instruction (Shohamy, 1992). Currently, with written grammar tests at the ELC, students prepare by studying their textbook and through written exercises. Students are not held accountable for orally producing the structures being tested. Therefore, an oral component included in grammar assessments may motivate students to attend more carefully to their grammar in speaking, thus helping to increase noticing, which will then lead to acquisition (see Doughty, 2003).

Many grammar textbook series, such as *Focus on Grammar,* include supplementary CDs to create worksheets, quizzes, and tests for written items and exercises. Other popular developments for teaching grammar are interactive software packages that allow learners to work autonomously through multi-media tutorials with animated grammar explanations (see http://www.elt.thompson.com and http://www.longman.com/ae/multimedia). These programs include assessment in the form of multiple-choice, fill-in-the-blank items, and matching, all of which are self-graded and often produce a progress report.

Some of these interactive programs have a speaking component, allowing learners to temporarily record their response and then listen to themselves and compare with a model. The new online grammar program, *Grammar Café*, by Thompson and Heinle, is one that actually provides a speaking component which learners are given speaking prompts to practice a particular structure (http://elt.thomson.com/namerica/en_us/technology/GrammarCafe.htm). The response is recorded and available for the teacher to listen and assess. However, the program is designed to be more student-centered than teacher-centered. Therefore, teachers cannot create or edit the tasks, nor do they have the flexibility of grouping several tasks together to get sufficient information about the students' mastery of the structures presently focused on in ELC grammar classes. Because of the limitations in the currently available software, it was necessary to develop a new program that could allow teachers at the ELC to create tasks that are appropriate for their students and select a group of tasks to effectively assess students' mastery.

Conclusion

This chapter has described the rationale for developing and implementing an oral grammar assessment by illustrating the grammar objectives and the teaching and testing methodology at the ELC. Because of the purposes of grammar instruction and assessment suggested by research, the current ELC teaching and testing procedures, and the limitations of current software programs for grammar, I decided to develop a separate oral grammar assessment to target those specific structures included in the course objectives for ELC grammar classes.

CHAPTER 4

PROJECT DEVELOPMENT

This chapter discusses the stages of development and implementation of the assessment tool I have created. The first section gives an overview of the program's basic format and design, including the objectives it was designed to meet. The second section describes how I developed the speaking tasks and the rubric. The third section outlines the technical process of developing the database and website to run the assessment program. The final section discusses the program's implementation and the changes I made during the first two semesters.

Project Overview

When I first conceptualized the development of an oral grammar test, I designed it as a diagnostic and achievement test to specifically assess the use of important grammatical structures covered in the grammar classes at the ELC. The test was to focus not on grammar knowledge, but on students' ability to use the grammar structures accurately and appropriately through speaking. Test scores would give teachers a specific measure of which grammar structures students have mastered, and which structures students have not mastered.

The format was a set of tasks in which students respond to a speaking prompt. The tasks could be organized into groups and administered under the discretion of individual teachers. From the beginning, I considered having responses recorded on the computer to be later scored by the teacher, similar to the ELC listening/speaking

classroom tests and LAT. I also considered the possibility of simply using face-to-face interviews with students, as well as in-class activities for assessment.

As I began developing the speaking tasks, my concept changed from just a test into a more complete assessment system and possibly an effective oral practice tool. I knew how to make websites and I knew early on that I wanted to involve the Internet. Initially, my programming skills were somewhat limited, and I had to learn new programming languages as I went along. As I learned more about what could be done, I had new ideas about how to build the program into a flexible system allowing teachers to interact with the website and create custom assessments.

The final product is an assessment system for the Levels 1-4 grammar classes at the ELC. It is designed to meet the following objectives:

- Allow teachers to listen to students' speech samples and measure mastery of specific grammar structures using a rubric
- Allow students to listen to their own recorded speech samples and assess their grammar using a rubric
- Provide speaking activities for teachers to use in class for students to practice using specific grammar structures
- Present a user-friendly, flexible system allowing teachers to create their own tests/speaking assignments and administer them either in class or as homework assignments

There are three main stages in using the program. First, teachers create a task selection. Teachers design their own test, called a "selection," by using an online database to select speaking tasks that will elicit grammar structures covered in their

course. They also have the option of writing their own tasks. This allows teachers to create the kind of assessment that best fits their syllabus. Teachers can also print out the tasks to use in class. The next step is for the students to go to the computer lab and complete the assessment**.** Students use the recording application, currently on all lab computers in the ELC, to log in, read the speaking prompts, and record their responses. Here, they can also listen to themselves and apply self- and peer-assessment. Their responses are saved in the online database on the server. They can do this during an in-class session or out of class. Finally, teachers can listen to their students' responses on the website and use the rubric to score students' grammar.

Developing the Tasks and Rubric

Selecting Appropriate Grammar Structures

In developing speaking tasks to elicit grammar structures, I began by carefully examining the course objectives for each level. The objectives are merely a list of grammar structures to be mastered at a standard of 80% on written tests. Level 1 has 21 structures; Level 2 has 34 structures; Level 3 has 28 structures; and Level 4 has 28 structures (see Appendix A for complete set of grammar objectives). I decided to exclude Level 5 from this project because the curriculum for these classes has transitioned to a content-area focus, rather than individual skill area classes. With the help of the ELC grammar coordinator, I identified the structures that are most distinguishing to each of the four levels, as well as being the most conducive to elicitation through a speaking task. I narrowed the structures down to about eight or nine for each level, to be a more manageable scope of material.

Writing the Tasks

The design for speaking tasks was simply to have a prompt that would elicit a specific grammar structure. Throughout this early development stage, I consulted regularly with the other grammar teachers, as well as using my own experiences teaching Levels 2 and 4. I asked the teachers what kinds of speaking activities they used to practice specific grammar structures. I also based some of the speaking tasks on written exercises in the *Focus on Grammar* textbooks (Fuchs & Bonner, 2006; Fuchs, Bonner & Westheimer, 2006; Schoenberg, 2006; Schoenberg & Maurer, 2006). Other tasks I based on the listening/speaking themes that have been integrated into the grammar courses. Pictures were often useful in eliciting the desired grammar structure. Originally, I had written approximately 2 tasks for each grammar structure. Figure 2 shows some examples of the tasks I wrote in the beginning stages of this project.

Level	Structure	Prompt
1	Modals "can/can't"	Look at the chart. Talk about what people can do and what they can't do. **CAN THEY . . . ?** / **AMANDA** / **JUDY** / **STEVE** / **JOSH** **SING** yes no no yes **DANCE** yes yes no no **SWIM** yes yes yes no **DIVE** no yes no no **WATER SKI** no yes no no **SKI** no no yes yes **PLAY THE GUITAR** yes no no yes **PLAY THE PIANO** no no yes yes
2	Comparatives	Think of two classmates. Compare their height, weight, personality, and skills.
3	Present Perfect with "already and yet"	Look at Jennifer's "To Do" list. Talk about what she's already done, and what she hasn't done yet. To DO: ~~grammar homework~~ ~~write 2nd draft of essay~~ ~~call my parents~~ ~~sign up for a writing tutor~~ pay health insurance ~~buy groceries~~ write an email to my cousin clean the kitchen
4.	Unreal Conditionals	Talk about a regret you have (something you did, that you wish you hadn't done, or something you wish you had done, but didn't). Talk about what you would have done differently.

Figure 2. Sample speaking tasks.

Throughout the next two semesters of implementing the system in ELC grammar classes, the task bank expanded and changed as teachers edited the tasks they used and developed some of their own. As the task bank grew, we also expanded the list of grammatical structures included in the assessment. Teachers often asked me to write tasks for a given structure that they were teaching at the moment.

Developing the Rubric

In considering what kind of rubric to use for these speaking tasks, I discovered that many rubrics used for measuring grammar ability in performance samples are holistic scales. For example, the new TOEFL speaking rubric contains a section on language use which holistically covers both accuracy and range of grammar structures (Educational Testing Service, 2007). The lowest level represents no control of any grammar structures. The highest level represents complete range of structures with no errors. This type of scale is not appropriate for this situation because we are not concerned with measuring their overall grammar ability, but rather the control of targeted features. Finally, I decided to try a 3-Level rubric for control over specific grammatical forms, similar to Purpura's model (2004, pp. 172-3). The rubric is shown in Figure 3. The lowest level represents no mastery of the targeted form. The middle level represents partial mastery, while the highest level represents complete mastery.

3	**Complete Mastery**: The structure is used correctly almost all of the time. There may be one or two mistakes in form, but overall the structure is used appropriately and with correct form most of the time. This shows that the structure has been mastered.
2	**Partial Mastery:** The structure is used correctly about half of the time. Sometimes it is used correctly, with the correct form. However, this accuracy is not consistent. Because the structure is used correctly only half of the time, it has been partially mastered.
1	**No Mastery:** The structure is used correctly almost never. There are many errors present in the form and way the structure is used. Because of these problems in accuracy, this structure has not been mastered.

Figure 3. The Rubric.

Piloting the Tasks and Rubric

During Winter Semester 2006, I taught two Grammar 2 classes. The other Grammar 2 teacher and I informally piloted some speaking tasks we had already developed. Throughout the semester, we had four written tests, and during three of those tests, we interviewed students individually in the hall. We gave them the prompt printed on a piece of paper. Students responded to the prompt and then I gave them a score and wrote it on their test. This experience helped me fine-tune the tasks because sometimes the wording was not very clear and the student asked for clarification.

Developing the Computer Program

The Database

I wanted the speaking tasks to be available through an online database, allowing teachers to sort by level and structure. MySQL is a system for managing online databases. Since MySQL is not designed for end users to directly interact with except through SQL commands, I used a web-based database administrator called phpMyAdmin (http://phpldapadmin.sourceforge.net/) to graphically create the database. An online database consists of tables to store information. Figure 4 shows the structure of the task table, with its fields for task identification number, level, topic, summary, preparation and

speaking time, grammar structure, and image information. As tasks are created, the information for each one is stored in this table.

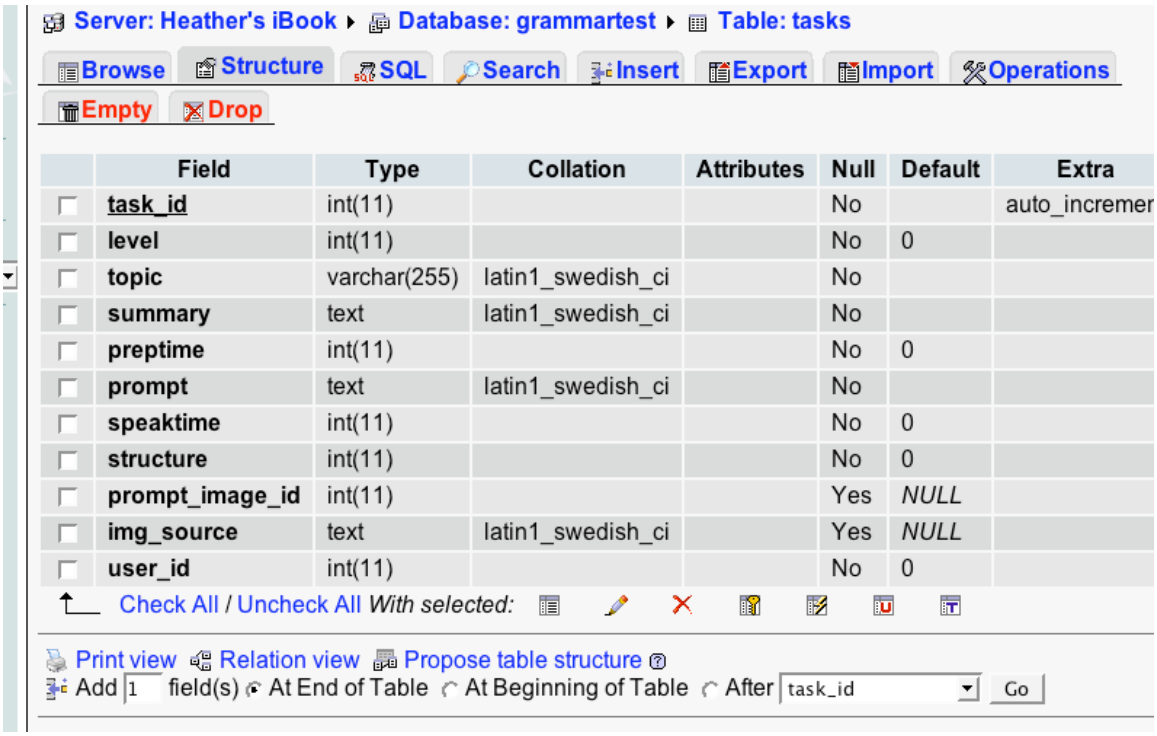

Figure 4. Structure of the task table within the online database.

I also made several other tables including a test selection table that is used for grouping the tasks into selections. To create separate user accounts, I made a user table with the following fields: name, username, and password. The complete database structure, including each table and fields, is shown in Appendix B. These tables could then reference each other as tasks are created, edited, and selected. With the database structure ready, I needed a web interface in order to actually display the information in the tables and interact with the data.

The Web Interface

The next step was to develop the web interface to access the database, allowing teachers to edit existing tasks, add new tasks, and create task selections. I chose to use PHP, a programming language that is commonly used with MySQL to build interactive web pages. Because I didn't know PHP initially, I had to learn it as I went. I bought a book to learn PHP on my own, and although the book was helpful, my brother taught me the most and coached me through the development phase. He taught me the needed PHP code as I determined what functions the program needed to perform. Fortunately, I found that the same code I made for one page could be copied and adapted to perform slightly different functions on another page. I created PHP pages to complete the following functions:

- Log into a user account
- View the complete task bank and sort by level and structure
- Add new tasks
- Edit existing tasks
- Create a new selection
- Add speaking tasks to the selection

In order to separate the PHP code and the HTML markup code, I used a library called the Smarty Template Engine (http://smarty.php.net). In this system, all of the markup that tells the browser how to display the web page is put in separate template files which are referenced by the PHP files. This makes the code clean and orderly. Creating the template files was easy for me because it used HTML, which I already knew. Also, I used a separate header template to display the menu bar and user information, which is

referenced by each template file. This way I did not have to recreate the header for each page. I then developed an external cascading style sheet to add a uniform style to the entire site.

By April 2006, the website was up and running. To enable me to easily update files, I chose to keep the website on my own hosted server. After looking at what was available, I finally registered the domain name grammaticallycorrectonline.com for the site. Since then, my grammar assessment program has been known as *Grammatically Correct.* The current URL is http://www.grammaticallycorrectonline.com.

The Recording Application

To create the recording component of the program, I used C++, a powerful programming language commonly used for writing standalone programs. I designed the recording application, which was coded by my brother. Because C++ is such a complex language, I did not have the time or resources to learn the language myself. In the beginning of May, 2006, the first beta version was downloaded onto all of the Macintosh computers in the ELC lab. During Summer semester, my brother was able to write an identical PC version which was downloaded onto the PCs at the ELC.

How Grammatically Correct Works

Creating the Assessment

Teachers log in to the website using an assigned username and password (see Figure 5). Then, after logging in they are taken through the task selection process.

Figure 5. Teacher login screen.

The process begins by asking for a name for the selection (see Figure 6). For example, if the selection were to assess past tense, it could be called "Quiz #2: Past Tense."

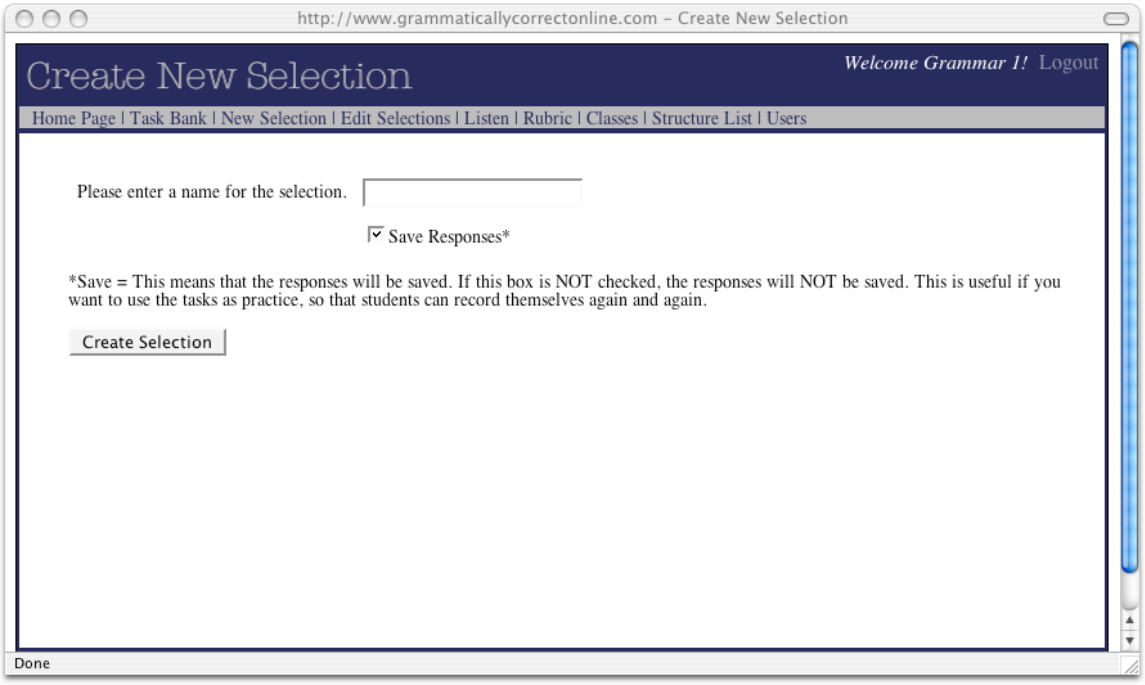

Figure 6. Making a new selection of tasks**.**

Then, teachers are taken to the task bank where they can select the tasks they want (see Figure 7). They can sort through the tasks by level and structure. They can also view the task details, including prompt images, by clicking the "View" link.

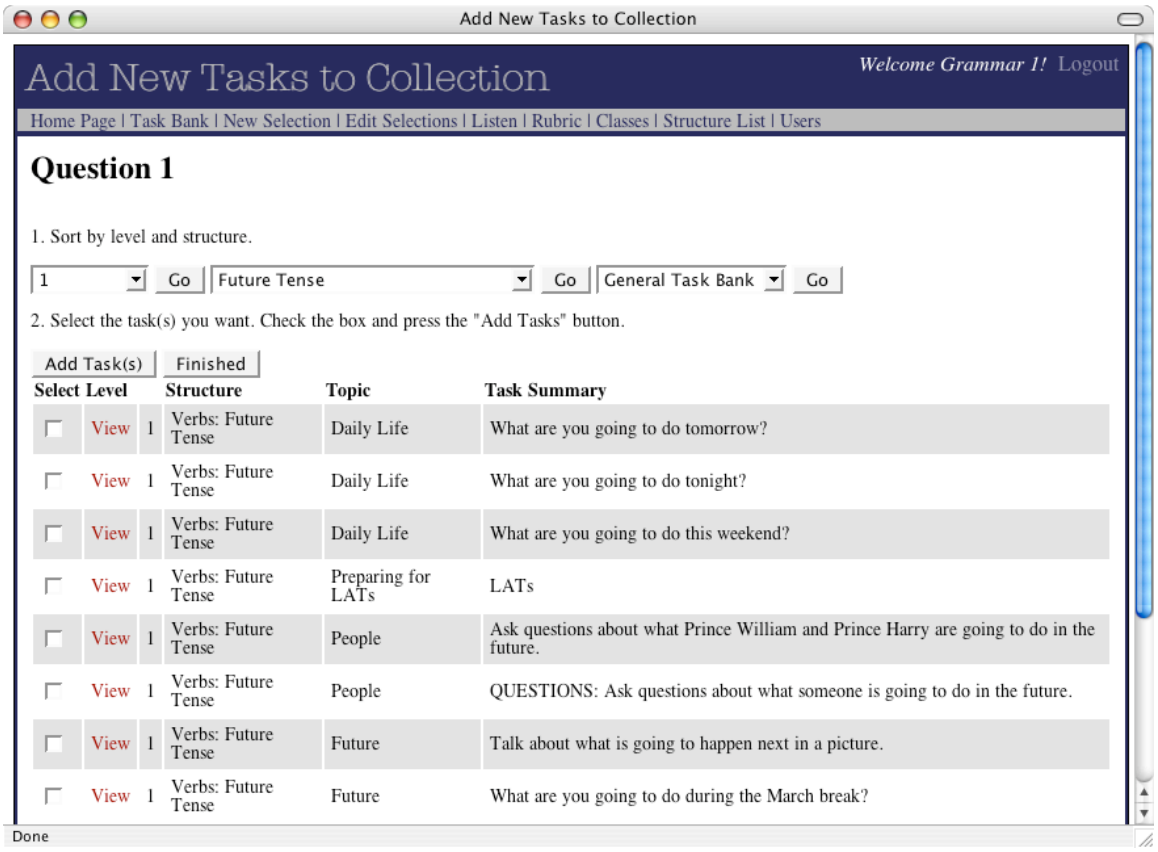

Figure 7. The task bank.

Teachers can also create their own tasks through the "custom task bank" where they can enter the level, structure, task summary, prompt, prompt images, preparation time, and speaking time (See Figure 8).

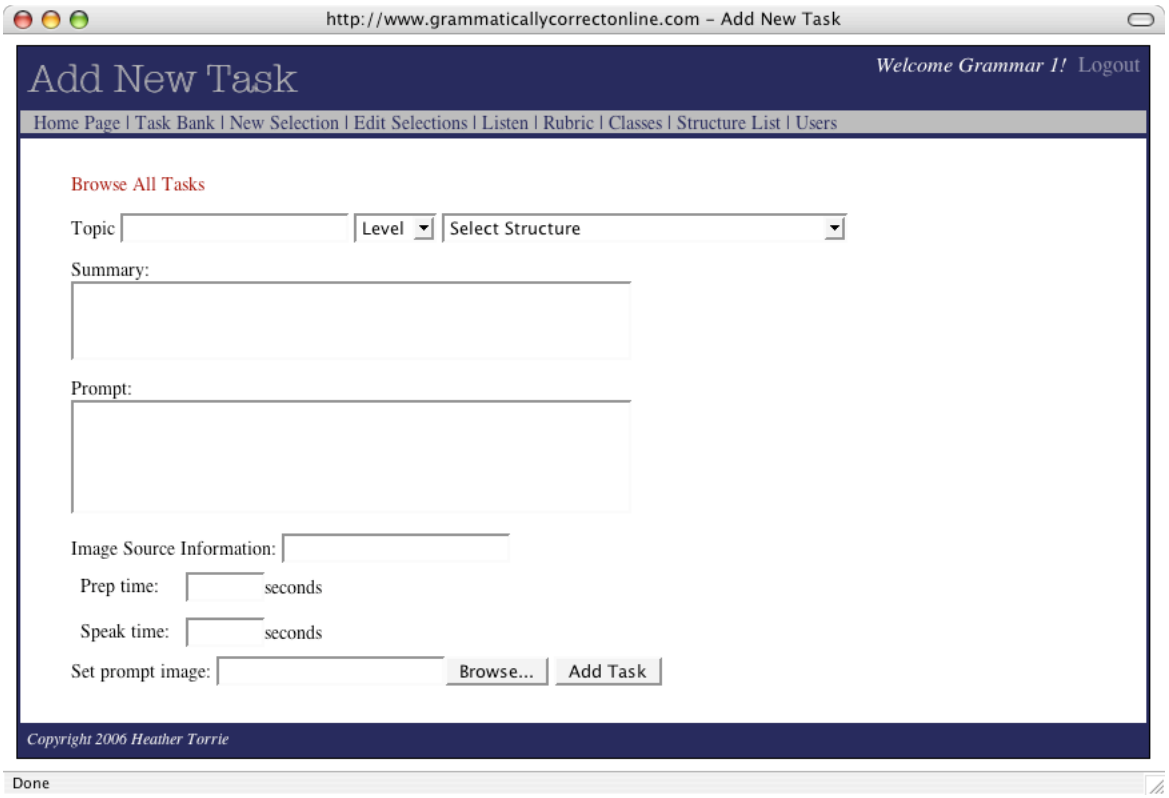

Figure 8. The form for making a new speaking task.

Now that the task selection is made, teachers can either print the tasks or have their students record their responses on the computer using the recording application. In order to have students access the tasks through the recording application, teachers must choose a start date and time and an end date and time. For example, a teacher might give students 2 or 3 days to complete the assignment. By clicking on the calendar icon, a pop-out window powered by Javascript enables teachers to easily choose the time and dates that the selection will be available to students for recording in computer lab (see Figure 9). The recording application can then communicate with the online database to display the prompts and record the audio responses.

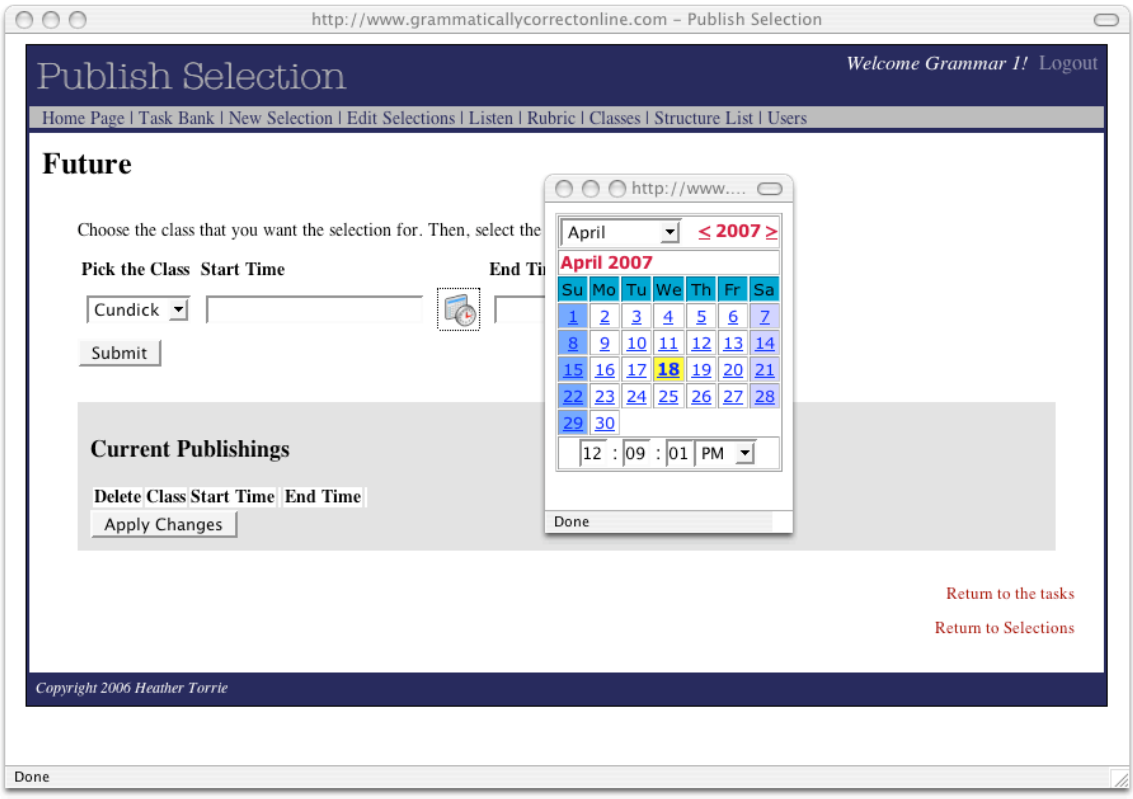

Figure 9. Setting start/end dates for the assignment.

Recording Task Responses

Students begin the assessment by opening the program with the *Grammatically Correct* icon. Figure 10 shows the opening screen, which has a microphone test and a log in button. My goal in designing the application was to have simple and clear instructions so students could easily test their microphone, log in, and begin with little coaching from their teacher.

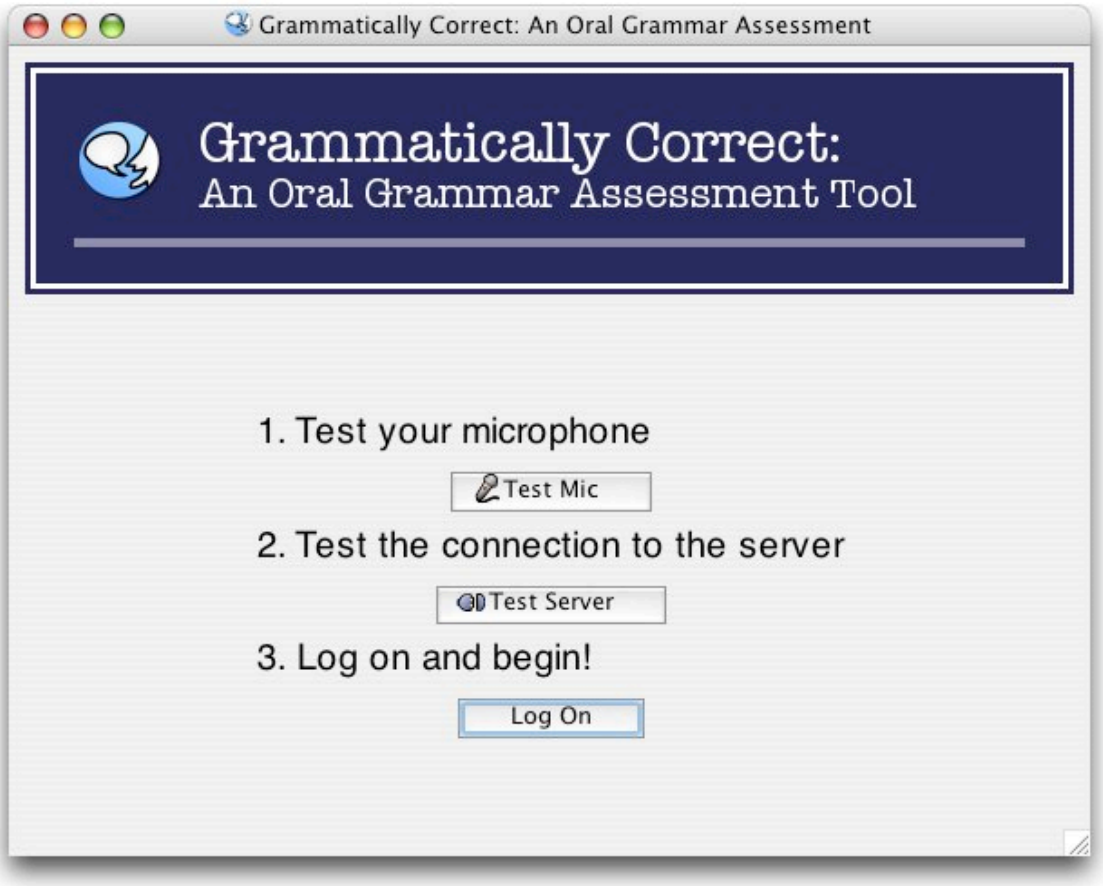

Figure 10. Opening screen of the recording application that students use.

Each task selection made on the website is connected to a particular class which has a password, created by the teacher. As shown in Figure 11, students log in by selecting their level, teacher, and assignment.

Figure 11. The login screen.

Students then type their name and their class password. This will then bring up the first task in the selection. Figure 12 shows how students are given an amount of preparation time in which they can read the prompt and think about what to say. As they prepare, the time counts down to zero.

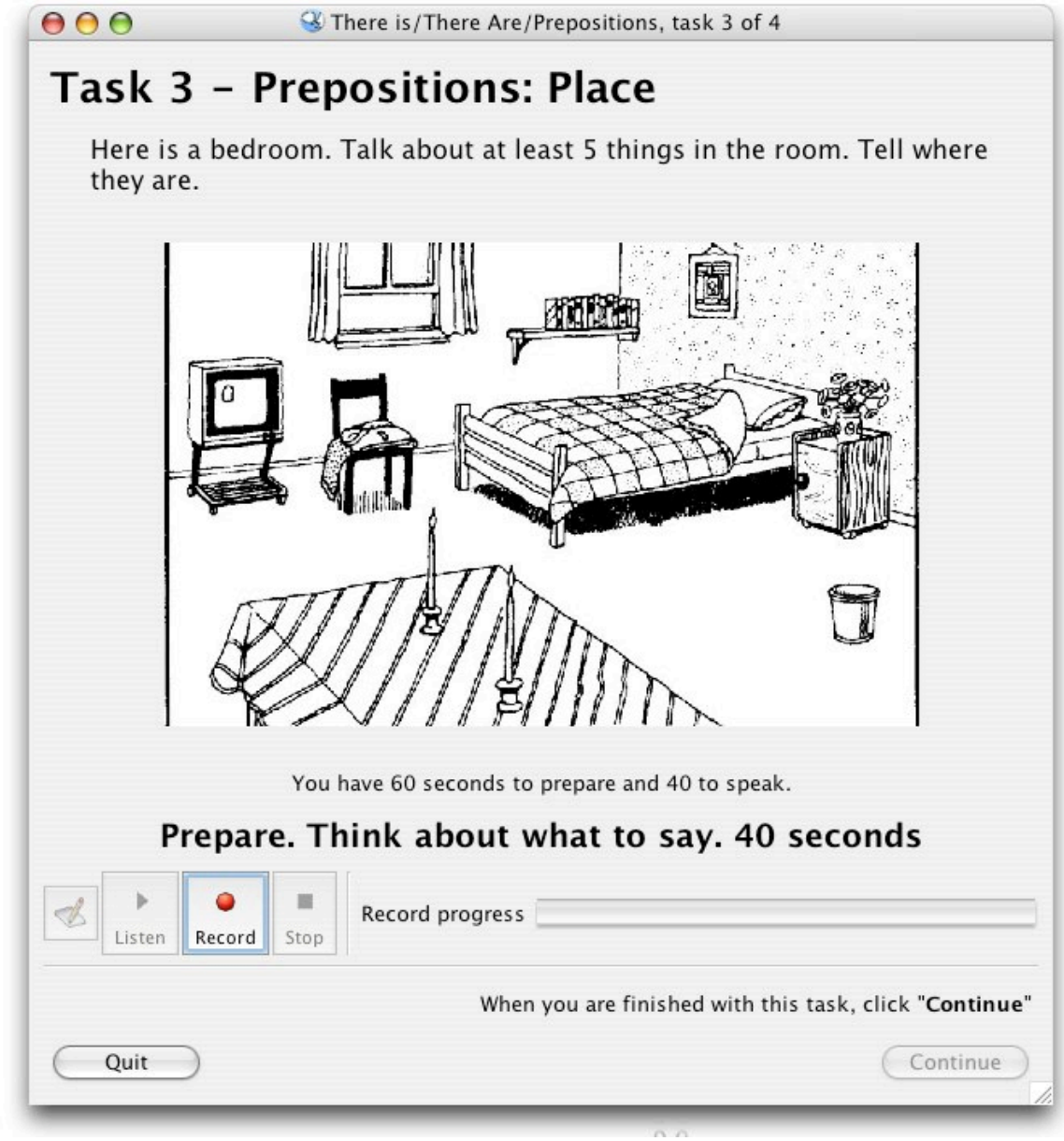

Figure 12. Sample Task: Preparation screen.

When the preparation time runs out, the program automatically begins recording (see Figure 13). However, students can also click the "Record" button if they are ready early. After recording, students can listen to themselves by clicking the "Listen" button.

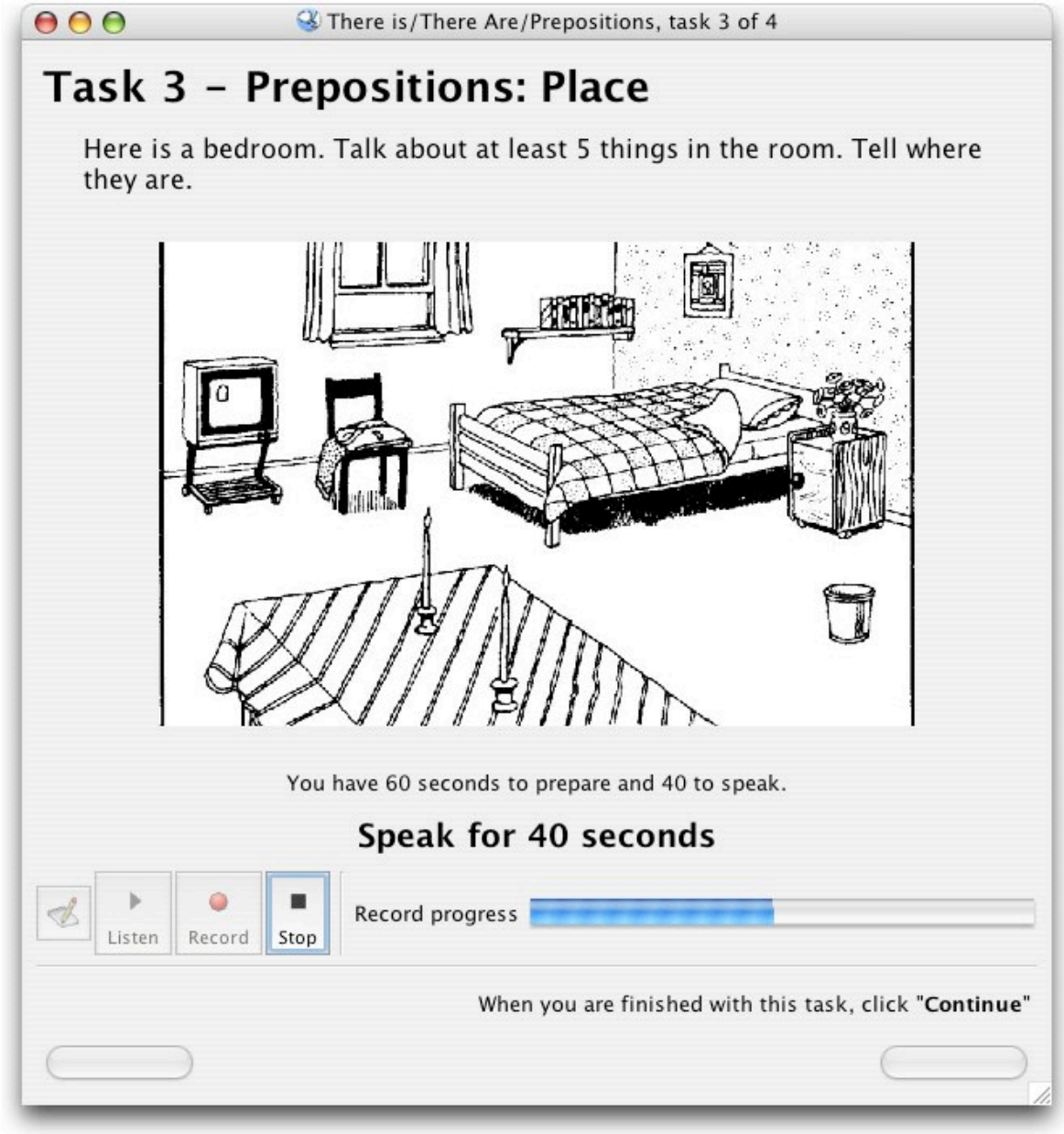

Figure 13. Sample Speaking Task: Recording screen.

Listening and Rating the Responses

After each completed task, the response is compressed into an SWF or a WAV file, depending on the settings, and saved in the online database. I set the default to use SWF, which are Macromedia Flash files, because they are smaller and take less time to

transmit to the server. Teachers can go to the website and access the sound files for each task selection (see Figure 14).

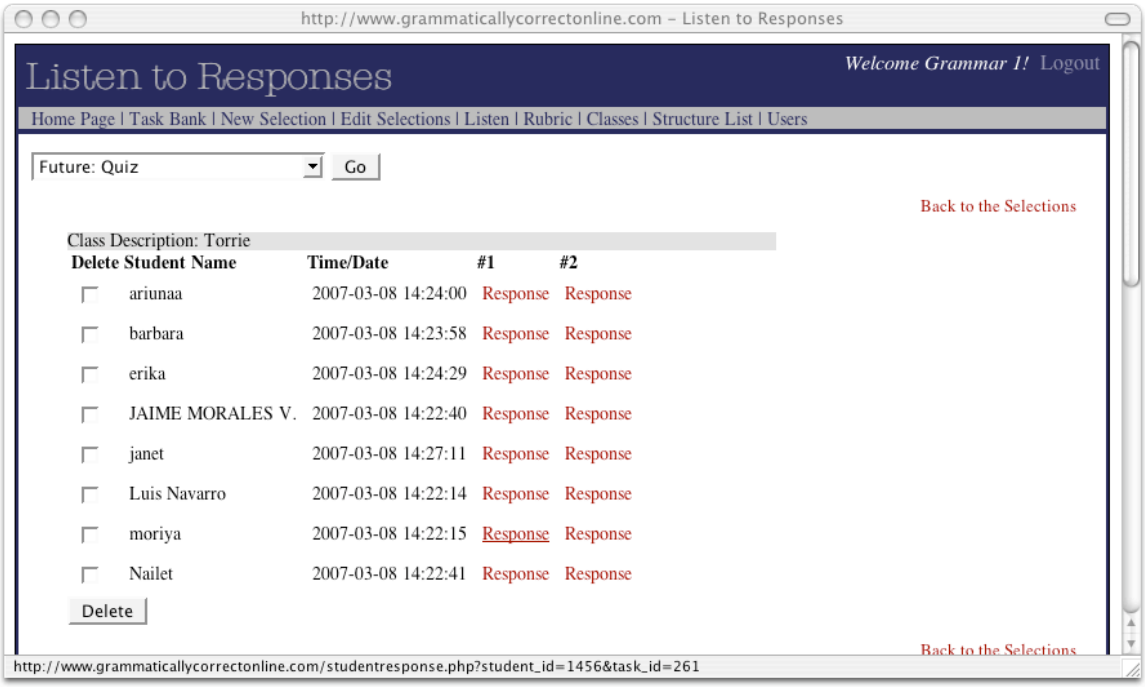

Figure 14. Links to students' responses on the website.

Implementation and Revisions

Summer Implementation and Changes

My objective for the first semester was to test the usability and technicalities of actually using *Grammatically Correct* with real grammar classes. We began the implementation process at the ELC in May, 2006, by meeting with all grammar teachers in a workshop. In the computer lab, I provided training through an instruction sheet and a demonstration of the website and recording application. Because I was unsure of how it should be used, I used the meeting to get teachers' input. The teachers divided into level groups and decided how to use *Grammatically Correct* in their class that semester. The Levels 1 and 2 teachers decided to try to use it in class and for a few homework

assignments. The Levels 3 and 4 teachers decided to use the program on review days before each test.

Throughout the summer, *Grammatically Correct* was mainly used during class time as practice. No formal scoring or feedback was given to the students, but a few of the Grammar 2 and 3 teachers gave their students a rubric to use during their in-class practice session. The students listened to their recordings and gave themselves a score, or they switched headphones with a partner and gave each other a score.

As we began using the program, further revisions to the website and recording application were necessary to make the program more user-friendly. The goal was to make the website self-explanatory and automatically take the teacher through the steps of creating a task selection. Most of the changes to the website involved simplifying directions and changing links. I worked to clarify and simplify the directions on the recording application as well. I wanted students to be able to navigate the program without looking at a printed instruction sheet.

A major challenge that arose early on was simply teachers' and students' unfamiliarity with the program. We were also experimenting with different ways to orient students. For example, some teachers used an LCD projector to walk them through the process of testing the microphone and logging in. Others had their students congregate around one computer and watch as they modeled how it worked. Some teachers wanted an instruction sheet for their students. However, reading through written instructions on a paper was not very effective either, as students wouldn't always read the instructions. Through my own experiences and talking to other teachers, I have found that the best way to train students is by having them congregate around the teacher's

computer and watch her walking through it. However, some teachers still like to give their students a handout.

Many of the revisions I made throughout the semester were based on teacher recommendations. For instance, one teacher suggested having a starting screen with instructions on it, rather than just automatically starting the timing process. The starting instruction screen is shown in Figure 15.

Figure 15. Starting instructions to students.

Another revision was to create a practice mode. Teachers who preferred to use *Grammatically Correct* as a practice activity during class time wanted their students to be able to record, listen to themselves, and re-record. To accommodate this, I created an option that teachers can select for practice mode. Under this setting, the audio responses are not actually saved onto the server, and the recording application allows students to record as many times as they wish.

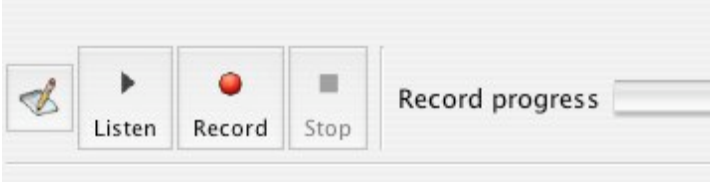

Figure 16. Revised buttons with the text-editor button added.

One of the teachers also suggested having a textbox for students to write notes after listening to themselves. To accommodate this, I added a button next to the "Listen" button that students can click to launch a text editor (see Figure 16). The other changes I made were minor, such as link names and the wording on website instructions.

Fall Implementation

During Fall semester, I wanted to involve more teachers and classes in using the program, as well as try having students complete assignments on their own outside of class. Unfortunately, we did not have the usual grammar workshop at the beginning of the semester, and I had to meet with teachers individually for an orientation. It was difficult to help teachers early in the semester because there were so many new grammar teachers and it was hard to schedule a time with each of them. Because of this slow start, the implementation and use of *Grammatically Correct* varied by teacher and level. In Levels 1 and 2 teachers mainly used out-of-class assignments, while teachers of Levels 3 and 4 used in-class practice sessions on their review days, as well as through homework assignments. Like the first semester we tried it, teachers did not use *Grammatically Correct* for actually assessing grammar. It was used as more of a practice tool. For the homework assignments, most teachers gave participation or homework points and not actual proficiency grades.

No major technical changes were made to the recording program or website during Fall Semester. Most of the bugs were worked out, and I focused on the evaluation.

CHAPTER 5

PROJECT EVALUATION

After completing the project development stage and initial implementation process, I then conducted a formal evaluation, applying Stufflebeams's (1996) guidelines of program evaluation in the design and procedure. This chapter discusses the project evaluation, beginning with the context, purposes, and evaluation questions. Then, the methods, including the participants, design, and procedure are discussed, followed by the results according to each evaluation question.

Context for Evaluation

As teachers used *Grammatically Correct* in their grammar classes during Summer and Fall semesters, I conducted some informal evaluation through class observations, talking to teachers, and asking students about their opinions and concerns. Also, many of the teachers initiated their own evaluation by asking students for feedback in the form of a discussion. The information gleaned from informal evaluation, discussed below, helped to establish the purposes and formal evaluation questions.

Informal Evaluation and Feedback

Teachers and students alike generally expressed positive feedback, saying that *Grammatically Correct* provided good practice for using the grammar structures in speaking. However, as Fall semester progressed, teachers also expressed frustration about using the program during class time. Although one of the original objectives was to allow students to use the program outside of class, many teachers still opted to use it in class to be sure students could get help if needed. However, they often ran out of time on

their review days and did not make it to the computer lab to use the program. Teachers of Levels 1 and 2 began giving out-of-class assignments, which helped save class time, but was also a concern in that teachers did not always take time to listen and give useful feedback to their students.

Another issue was that of creating the task selections. The program is designed to allow teachers the flexibility to make their own custom speaking assignments, which can fit into their regular course calendar. However, some teachers expressed the feeling that they would benefit from having the task selections already set up, rather than having the flexibility of creating their own selections. Furthermore, some teachers did not use the program at all because they were either unsure of how to use it in their classes, or they simply were not willing to invest the time and energy into implementing a new program.

Other informal conversations with students and teachers related feelings of frustration with technical problems, including problems with microphones and the lack of computers for every student during in-class lab sessions. Also, some students didn't like the in-class sessions because of the noise when all students are talking simultaneously. Despite these issues, most students said they liked using the program and it was a helpful tool.

Purpose of this Evaluation

Based on informal evaluation, the program appears to have some perceived value. However, a formal evaluation was needed to solidify its perceived effectiveness, and to clarify concerns, as well as to provide information about how the program should be implemented at the ELC. I also expected that improvements to the task development

system and usability would naturally be suggested, allowing me to make any necessary changes.

Because the program has not been fully implemented, this evaluation was of a more limited scope. At this early point, it was important to determine whether teachers and students felt that it is effective enough to continue implementing. If they perceived it to be effective, the program could then be evaluated on a larger scale to look at other issues, such as actual assessment of grammar proficiency, reliability, and validity of using its test scores.

Evaluation Questions

Based on the purposes established above, three main evaluation questions were formulated. The first question concerns the effectiveness of *Grammatically Correct*, with the criteria used being whether teachers and students—the two main groups of stakeholders—consider the major areas of the program to be useful and effective. If teachers and students perceive it to be an effective tool, this evaluation will also answer the second evaluation question of how to implement the program. It is clear from the experiences of Summer and Fall semesters that a standardized implementation plan is needed to help new teachers know how and when to use this assessment tool. The three evaluation questions are summarized here:

1. How effective is *Grammatically Correct?*
    a. Do teachers feel that the tasks and rubric are effective, as measured by the potential for assessment and the level of benefit to using the program?
    b. Do students feel the program provides meaningful practice and assessment?

c. Do students like using it?

2. How do teachers and students feel the program should be implemented at the ELC?
3. What improvements should be made to the program?

Methods

Participants

Information was collected from the two main groups of users--teachers and students. Table 1 shows the number of teachers, classes, and students enrolled at the ELC in Levels 1-4 during Fall Semester 2006, along with the number of those who participated in the evaluation.

Table 1

Number of participants in the evaluation

	Total Teachers	Participating Teachers	Total Classes	Participating Classes	Total Students	Participating Students
Level 1	2	1	2	2	18	18
Level 2	2	1	3	3	33	33
Level 3	4	4	5	5	82	82
Level 4	4	1	7	2	108	34
Total	12	7 (58%)	17	12 (71%)	241	167 (69%)

All Levels 1-3 teachers and classes used *Grammatically Correct* during the semester and participated in the evaluation, except for me. I taught one Grammar 1 class and one Grammar 2 class, and therefore excluded myself, but not my students. As for Level 4, only one teacher used the program and participated in the evaluation, along with his two classes.

Design

To provide triangulation of data sources, information was collected in the form of surveys, followed up with class discussions and interviews with teachers. The items on the surveys acted as a springboard to more fully answer the evaluation questions during discussions and interviews.

The development and validation of the surveys took place during Summer and early Fall semesters, 2006. Conversations with teachers helped in further developing the information needs and formulating the items included on the surveys. Changes to the surveys included adding more items concerning overall effectiveness and suggested use, and deleting items that dealt with individual tasks. To see that the items adequately addressed the evaluation questions, my committee chair, as well as another grammar teacher, helped examine the revised surveys and gave feedback. Furthermore, because the students are English-as-a-second-language learners, the survey for them needed to be understandable to all four proficiency levels at the ELC. In order to ensure readability, as well as validity, two students went through the student survey and gave feedback on the clarity of the instructions and the items.

Procedure

Surveys. The development and validation of instruments was finished in October, in order for the IRB approval to be processed in the beginning of October. The surveys for teachers consisted four sections. The first section asked teachers to rate the effectiveness of the program for practice, and then for actual assessment purposes. Two open-ended questions asked what they liked most about the program and what their most important concern was. There were also several likert items and open-ended questions

regarding the effectiveness and clarity of the tasks and the rubric. The second section of the survey was about implementation and asked teachers to report on the ways they had used the program, and then how they thought it should be used in the future for their respective level. The third section simply contained open-ended questions for suggestions on improving the website and recording application. A final section allowed room for any additional comments about the program. The complete teacher survey can be found in Appendix C.

The student survey consisted of ten likert items with statements regarding issues of importance to students, including usability, clarity of directions, topics, preparation and speaking time, self-assessment, and feedback. After the likert items, students were asked how often they wanted to use the program and in which ways. The last section on the student survey contained two open-ended questions asking what students liked and what they did not like about the program. The complete student survey is included in Appendix D.

During the second week of November, I distributed the surveys to seven grammar teachers who used *Grammatically Correct*, allowing them adequate time to complete the surveys before the end-of-semester activities. Then, during the first week of December, surveys for students were distributed to the twelve classes that used the program that semester. For Levels 1-3, students were given one last *Grammatically Correct* homework assignment that week, and the surveys were given out at the same time so that the program was fresh in their minds. To decrease bias in the procedure, the surveys were distributed by the teachers, rather than by me. In this way, hopefully students were more open and honest in their answers. The surveys were then collected at the end of the

week, and teachers were encouraged to give participation points for returning the surveys. All of the seven teacher surveys were returned, and of the 167 student surveys given out, 115 were returned. Table 2 summarizes the total number of students and surveys returned by level. A few of the surveys were not included because either the student did not sign the consent form or the student did not answer every question, which can be expected in a pencil-paper survey.

Table 2

Number of surveys distributed and returned

	Level 1	Level 2	Level 3	Level 4	Total
Surveys Distributed	18	33	85	34	167
Surveys Returned	16	30	54	25	115
Percent Returned	89%	94%	64%	74%	69%

Class Discussions. Because of practicalities such as scheduling, I held one class discussion in each level. These discussions were held the last week of classes, after students had completed the survey. I went into the classroom for about 10-15 minutes and tried to conduct the discussion as objectively as possible, not explicitly stating that I was the developer. I first reminded them of the computer program by showing the opening screen on the computer. Most students remembered using it that week and at other times throughout the semester. I asked general questions regarding issues from the surveys, including what they liked about the program, what things could be improved, and how it should be used for their level. Questions varied slightly from class to class depending on students' comments. For instance, when students reported not liking something, I asked specific questions to clarify the concerns. I took notes of the comments and how many students gave similar comments or expressed agreement.

Teacher Interviews. I held interviews with four grammar teachers, one from each level, in December and January. Each interview was approximately 20 minutes long, in which I asked open-ended questions regarding their feelings towards *Grammatically Correct.* Among the questions I asked were what they liked most about the program, and what concerns they had. I was particularly interested in how they thought it should be implemented in their level for future semesters, and what changes they recommended.

Data Analysis

The items on the teacher and student surveys were categorized into groups according to the criteria under the evaluation questions they answered. Then, each item was analyzed according to its quantitative or qualitative nature. Information gleaned from the class discussions and interviews are included in the qualitative analysis section.

Quantitative Analysis. Below are listed the types of quantitative items included in the survey and the analysis procedure I used for each one.

- Multiple Choice Items: To present a clear distribution of which options reflected students' and teachers' position most strongly, I calculated the percentage of responses to each option. For the teachers, because there are only seven, I used numbers only.
- Likert Scale Items: As with the multiple choice items, I calculated the percentage of responses to each option on the likert scale. This provided a clear interpretation of the feelings of students and teachers regarding the items. Using percentages also helps show the variation without having to calculate means and standard deviations.

- Checkbox Items: I calculated the frequency of each option and displayed the results in a chart for each item.

Qualitative Analysis. For the qualitative items, I transcribed and categorized the comments into thematic groups, allowing new themes that arose to be taken into account. Comments from class discussions and interviews are discussed according to common themes, interwoven with quantitative data in the Results section.

Results

The results of this evaluation have been summarized according to the main evaluation questions. To illustrate perceptions of teachers and students, I have included charts summarizing the results of the relevant quantitative items on the surveys, as well as qualitative data in the form of quotations from students and teachers to support the quantitative results. Quotations are from comments written on the surveys by students and teachers, and also from class discussions and interviews.

Question 1a: How effective is Grammatically Correct, according to teachers?

This first question regards teachers' perceptions of the program, specifically the tasks, rubric, potential for assessment, and overall benefit; the results are presented below according to those categories.

Effectiveness of the speaking tasks. For the most part, teachers agreed that the tasks were effective in eliciting the desired grammar (see Figure 17).

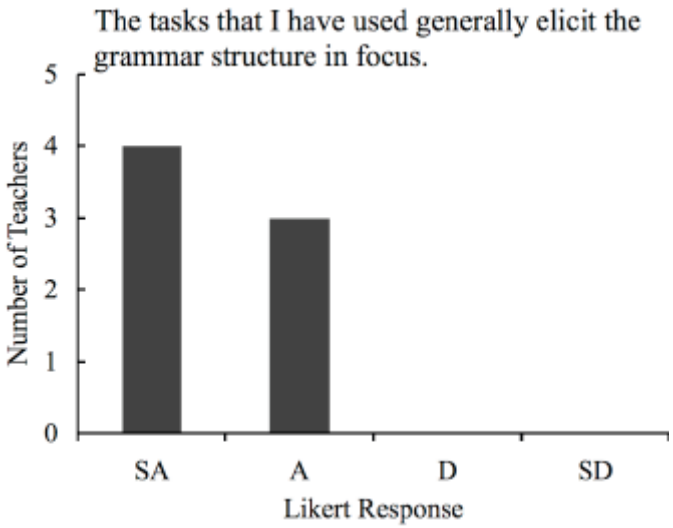

Figure 17. Teacher feedback on the effectiveness of the tasks in eliciting the desired grammar.

Qualitatively, however, the evaluation also produced some areas of weakness in the tasks, including readability of the prompts and time allotments for preparation and speaking. These concerns were brought up briefly in comments from the teachers, but were even more prevalent in the student feedback, and therefore they will be discussed with the student feedback.

Effectiveness of the rubric. In examining teachers' perceptions of the effectiveness of the rubric, we can first look at how it was actually used during the semester. Figure 18 shows the different ways the rubric was used this semester. It is evident that none of the teachers used the rubric to give feedback to students. Only one of the teachers reported actually using the rubric, and she used it as self- and peer-assessment during in-class practice sessions, while the others did not use the rubric at all. Two teachers did give feedback to their students, but in another form.

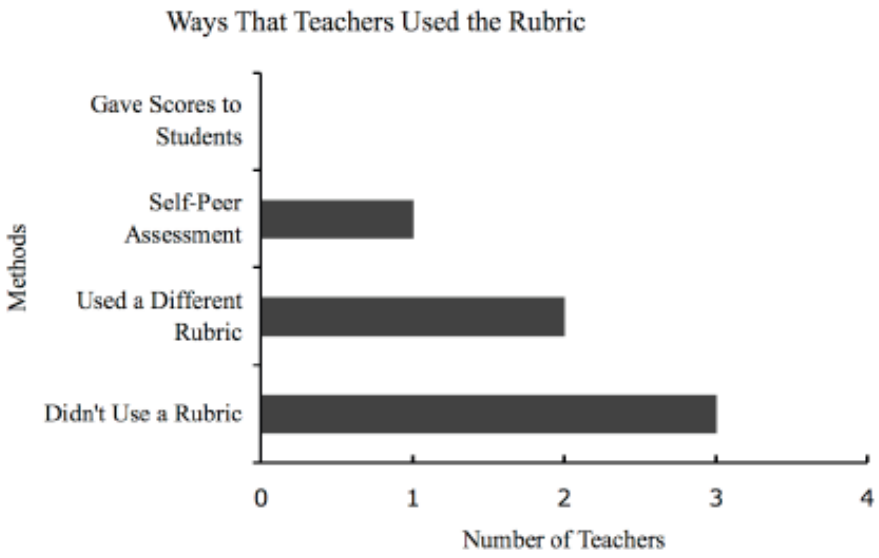

Figure 18. Teacher feedback on how the rubric was used.

To explain her reason for not using the rubric, one teacher wrote, "It was just laziness. I listened to their responses and gave feedback, but it was directly applied to what they said and the grammar we studied. I wanted their feedback to apply to the problems I heard." In an interview, another teacher explained that he liked to listen to all the students and look for common problems. Then, he would address the problems he heard in class.

Despite the lack of actual use of the rubric, most of the teachers agreed that the rubric could be effective in helping teachers determine their students' mastery of grammatical structures (see Figure 19).

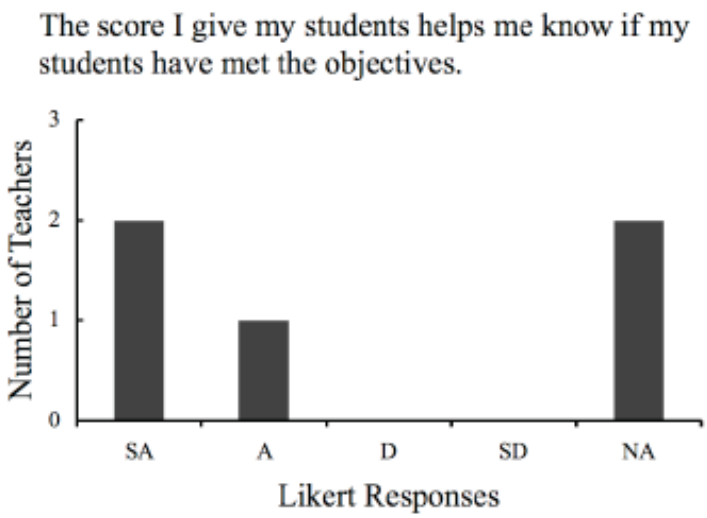

Figure 19. Teacher feedback on the effectiveness of the rubric.

The teachers also agreed that the rubric descriptions were clear (see Figure 20).

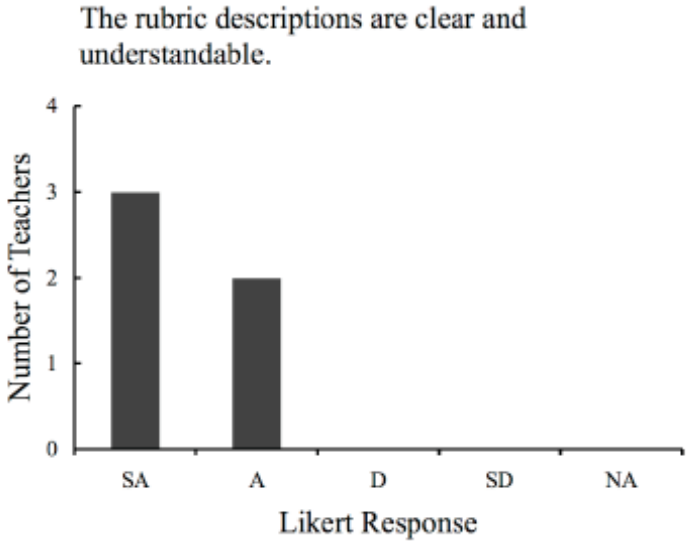

Figure 20. Teacher feedback on the clarity of the rubric.

However, there is some question as to the usefulness of scores to the students. Although all teachers agree that the rubric is clear and the levels are distinguishable, one concern is how useful the rubric is in giving effective feedback to the students. A teacher wrote, "Might be effective in teachers scoring a test, but I'm not sure how effective it is to give

feedback to a student after time has passed (few days later, weeks, etc)." Other comments were related to the score being overly simple and students wanting feedback in more detail. Even with a score, some teachers feel that students need immediate corrective feedback to avoid repeating mistakes. In fact, one teacher said that because of the lack of corrective feedback as students are using the program, she prefers "class activities which provide more 'real' speaking practice and peer/teacher interactions."

There was also a concern about using the rubric when the tasks are sometimes so short that they only elicit the structure a few times. To explain, a teacher wrote, "Since there isn't too much time allotted for each task, when students speak, they only use some structures about 3-4 times. In complete mastery, 1-2 mistakes would only be 50% accuracy, right?" One teacher also said that he preferred to use a scale with four levels: two high scores and two low scores.

The potential for assessment. On the survey, most teachers reported that the program was effective as both a practice tool and for actual assessment (see Figure 21)

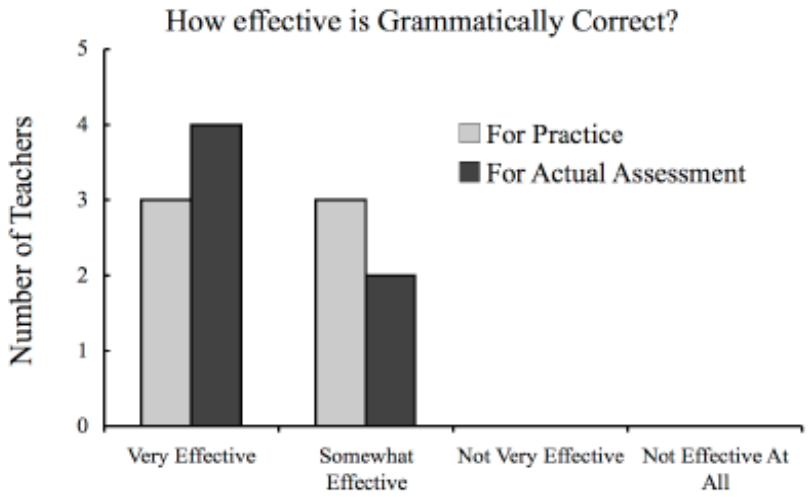

Figure 21. Teacher feedback on the program's effectiveness for practice and actual assessment.

However, in the classroom so far, teachers have been using the program primarily for practice, and not assessment. I tried to clarify this issue through the teacher interviews, and found that teachers' confidence in using the program was a major factor in determining how it was implemented. One teacher explained, "I think that if the teachers were confident in using [*Grammatically Correct*] and told the students from the beginning that there will be an oral evaluation of their grammar, it would work well as an assessment."

Besides the lack of confidence in using the program, another challenge seems to be the time requirement to actually use it for assessment. When asked what their most important concern was with *Grammatically Correct*, one teacher said, "not having time to listen to student responses and having students expect that I will." Another said, "Time needed to make it effective for students and for teachers." This reflects many other informal conversations and comments from teachers before this evaluation and shows the challenge of implementing the assessment in a way to minimize the burden on teachers and fitting it into the course content. This comment, given in an interview, summarizes many of the smaller comments from teacher surveys: "I think if I had or made more time, I would use it as a testing tool, but it hasn't been done before so it is a new concept for many teachers and of course you want to stick to the easy written tests that are quick to grade."

Overall Benefit. Despite the concerns with using the program for actual assessment most teachers seem to think it is beneficial. Teachers commented that the program is helpful in integrating L/S and making the grammar less abstract. It "[m]akes grammar real and takes it out of a book context," one teacher said. Teachers also liked

that students can do it in class or outside of class, and that all students could practice simultaneously.

Summary of Teacher Feedback. Overall, teachers believe the tasks and rubric are effective; however, some problems with task planning and speaking time, as well as prompt readability were brought up by teachers and further clarified by student feedback. As for the rubric, some teachers feel that scores do not give sufficiently detailed feedback to students. Teachers also feel that the program has the potential to be used for assessment, but they often lack confidence to use it for more than just practice. Some teachers also feel that to really use it as an assessment tool, it would take more time and energy than they have. Looking beyond assessment, most teachers think the program is beneficial in making grammar less abstract and providing practice opportunities.

Question 1b: How effective is Grammatically Correct, according to students?

Turning now to student feedback, their perceptions of effectiveness are categorized according to the themes that arose through the evaluation: practice, self-assessment, and general attitudes towards using the program.

Practice. Of the students participating in the evaluation, 86% felt that the program is useful in helping them practice the grammar they learn (see Figure 22).

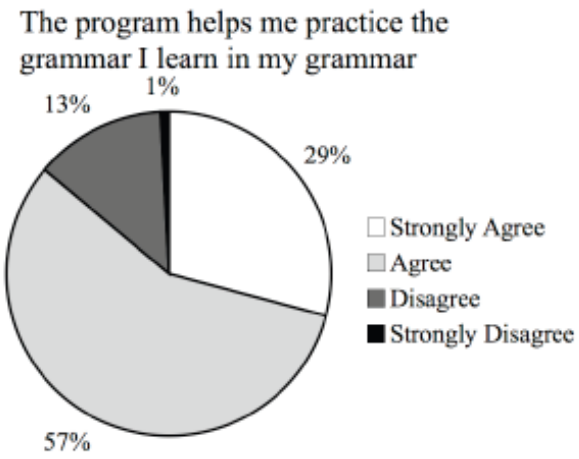

Figure 22. Student feedback on the benefit of the program for practice.

Qualitative data also lends support in this area, and actually the largest portion of the comments from students is centered around the benefit of practice. In answer to what they liked most about the program, there were 27 similar comments regarding the benefit of practice. One student wrote, "The combination to talk and grammar because I can to practice good grammar," while another student said "We can practice about grammar when we learn in grammar class. It is good for homework." Similarly, another said, "It is a lot of helpful this program because most the time the grammar that we learn we can't use it in speaking it's hard do it, but if we practice that will be more easy after." More comments like these were made in the class discussions. For example, two students in Level 4 said it was good practice, and several nodded in agreement. Three students in Level 1 and three students in Level 2 said that it was good to practice what they learn in class.

Self-assessment. The majority of students said that listening to themselves was beneficial. They also seemed to believe that they could hear when they made a grammar mistake (see Figure 23).

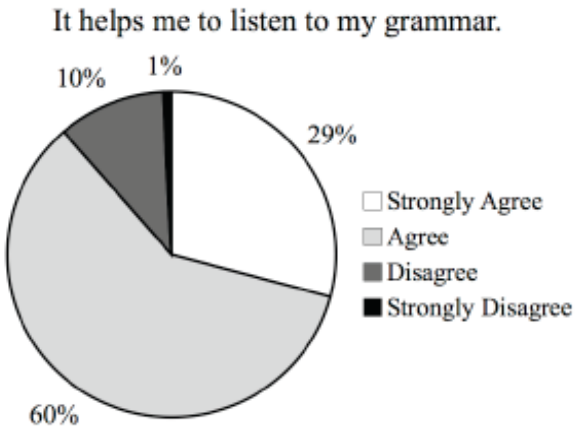

Figure 23. Student feedback on the benefit of self-assessment through listening to themselves.

Another student said, "It's helps me to know how I am speaking clear and I can hear where I made a mistake and gives me practice (real) for my grammar." This reflects 19 other comments along the same theme, such as "I can listen my mistake" and "I find that mistake, when I spoke."

Question 1c: Do students like using the program?

Most of the students, 88%, said they liked using the program (see Figure 24).

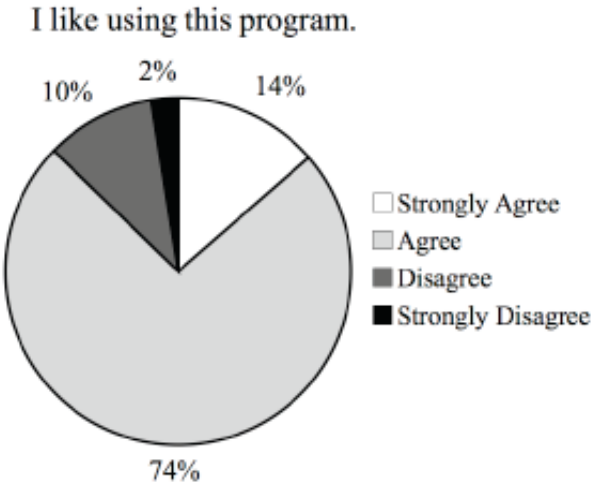

Figure 24. Students' overall opinion on using the program.

Qualitatively, student feedback from the surveys can be categorized into three general areas. First, 13 students simply said they liked it. Second, 6 students felt it improved their English. A student said, “it can help me to improve my speaking.” Third, 11 students said they felt that the program helped them to actually use the grammar they learn in grammar class. “Even though we learn Grammar, we often don't use this Grammar. But this program. It'll improve my English.”

Naturally, not all students liked using *Grammatically Correct*. Three students wrote that they did not like the program. "I don't feel very useful," one of them said. Another wrote, "I prefer grammar quizzes to it." The specific concerns brought up by students were mainly of two categories. First, several students said they didn't like speaking into a computer. Four students commented on the survey that they feel nervous when they are in front of a computer, and it is not the same as talking to a real person. In the survey, one student wrote, "Maybe the thing that don't like me its that sometimes you can make nervous to talk in a computer and maybe you can do it great with some people but do it wrong in the computer..." Two Level 2 students echoed this same notion in their class discussion.

The second main concern students expressed was the need for feedback. Throughout the semester, not all teachers listened to their students and gave any feedback. On the surveys, six students specifically said they needed feedback. Two sample comments say, "I dislike no check-my recording voice," and "I don't know my score." This concern was affirmed in class discussions. In each class that I visited, the majority of students said they wanted their teacher to listen to them and give them feedback, and the students who said that the program wasn’t helpful explained that the

reason was because of lack of feedback. A Level 4 student explained that it was not enough to just get a score. She wanted immediate feedback, such as that in speech recognition software, to help her correct her grammar mistakes.

Tasks. Three main concerns brought up by the students are the readability of the prompts, task topics, and preparation and speaking time. First, the main problem with readability is that the picture is sometimes too small to see. For example, one student wrote, "Some pictures I can't see." I have been aware of this problem, and it is unique to specific tasks that we developed.

Secondly, five students commented that some of the topics are boring or confusing. For example, one student wrote "Sometimes topics were confusing." Another student wrote, "It depends on the topic. Some subject is really boring."

Finally, the most widespread complaint was with the allotted preparation and speaking time. As illustrated in Figures 25 and 26, almost one third of the students disagreed with the statements "I have enough time to plan" and "I have enough time to speak."

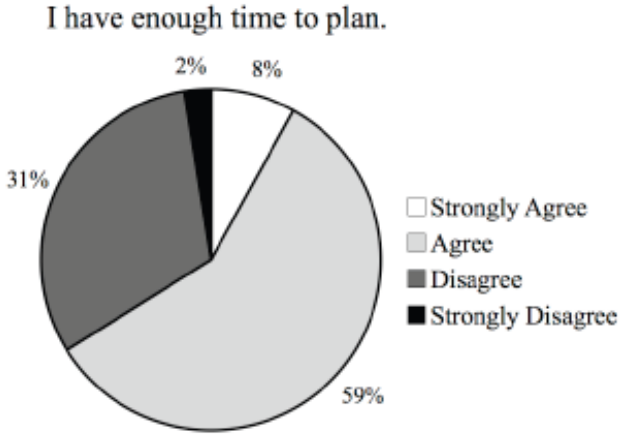

Figure 25. Student feedback on the amount of preparation time.

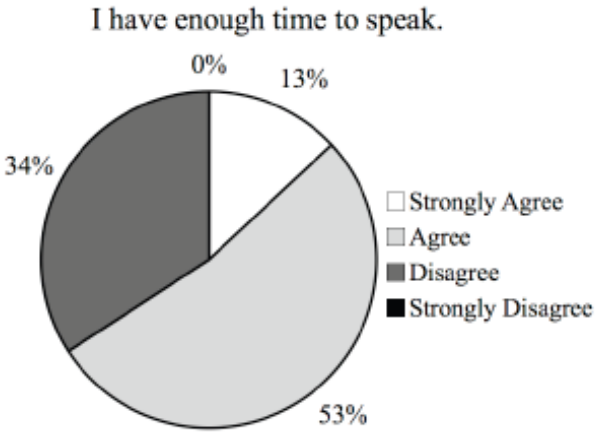

Figure 26. Student feedback on the amount of speaking time.

This concern was also reinforced by comments from students and teachers. In fact, the majority of the student comments on the survey were related to not having enough time to plan and speak. As examples of the 28 student comments on this issue, one student said, "We don't have enough time to think what we'll speak about the questions," while others simply said, "Not enough time" or "Sometimes it is fast." Another student wrote, "We don't have enough time to finish the speaking." As I held the four class discussions, I noticed that this was also one of the most prevalent issues brought up. When I asked students what they didn't like about the program, one of the first comments usually related to preparation and speaking time, especially in Levels 1-3. About 10 students in those classes explicitly said they wanted more time. However, this concern was not brought up by the Level 4 students

Summary of Student Feedback. Students seemed to like the program, mainly because of the opportunity to practice the grammar in speaking, and for self-assessment as they listen to themselves. However, two primary concerns were anxiety when speaking into a computer and the need for feedback.

Question 2: How do teachers and students feel the program should be implemented?

During Fall semester, grammar teachers at the ELC reported to have used *Grammatically Correct* 3-5 times for each class for in-class practice and also as out-of-class homework assignments. To gather suggestions for future implementation, I have summarized comments from the teacher survey and interviews, along with student feedback. On the student survey, one item was a checkbox item asking students to check all the ways they would like to use the program in their level, while another item asked students to choose how often they would like to use it in their level. Summarized below is the information gathered from the surveys, class discussions, and teacher interviews regarding implementation for each level.

Level 1. For Level 1, the teacher said once every two weeks or for each test would be the best frequency of use. Also, she thought that in-class practice sessions are best to orient the students to the program. Once they are familiar with it, the program should be used other weeks as an oral component with each testing unit, as well as for some homework assignments. Figures 27 and 28 represent student feedback, showing that ten students said they wanted in-class practice, six students wanted to use the program after class, while only two students wanted to use it on each grammar test. As for frequency, 50% of all the Grammar 1 students said they wanted to use the program every week, and 44% wanted to use it every day.

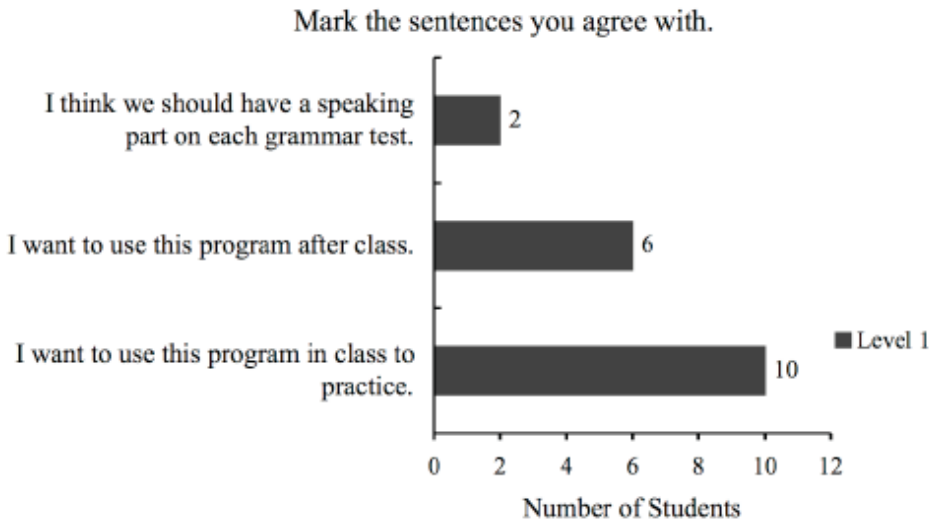

Figure 27. Student suggestions for methods of implementation for Level 1.

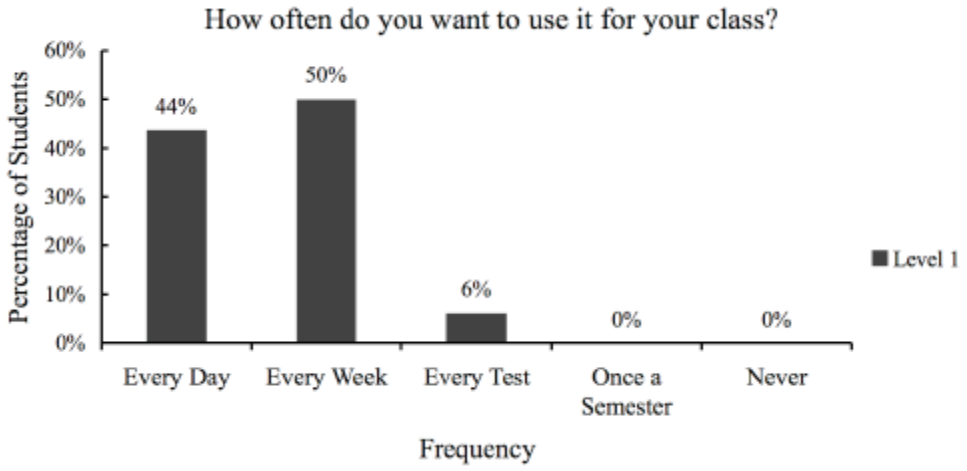

Figure 28. Student suggestions for frequency of use for Level 1.

Level 2. The Grammar 2 teacher thought that a combination of in-class practice sessions and out-of-class assignments scored by the teacher would be best for Level 2. In an interview, she said that the next semester, she wanted to try to actually use the program for assessment and keep track of students' progress. She said it should be used every 3-4 weeks, depending on the units taught. Student feedback is shown in Figures 29 and 30. Of the Grammar 2 students, 14 said that they wanted to use the program in-class

to practice, and 14 wanted to have out-of-class homework assignments. 11 students said they wanted to have a speaking part on each grammar test. Using the program for actual assessment on each test was also a popular choice when students were asked how often they should use it. 37% of the Grammar 2 students said they should use it for each testing unit. 23% wanted to use it once a week, while 33% wanted to use it every day.

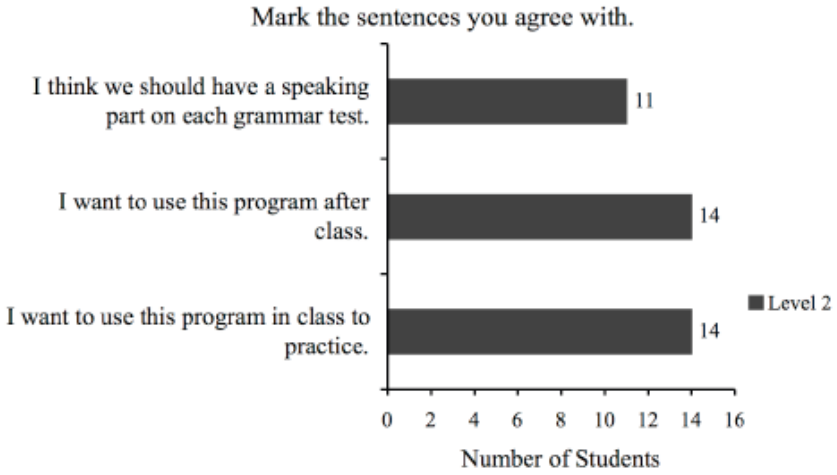

Figure 29. Student suggestions for methods of implementation for Level 2.

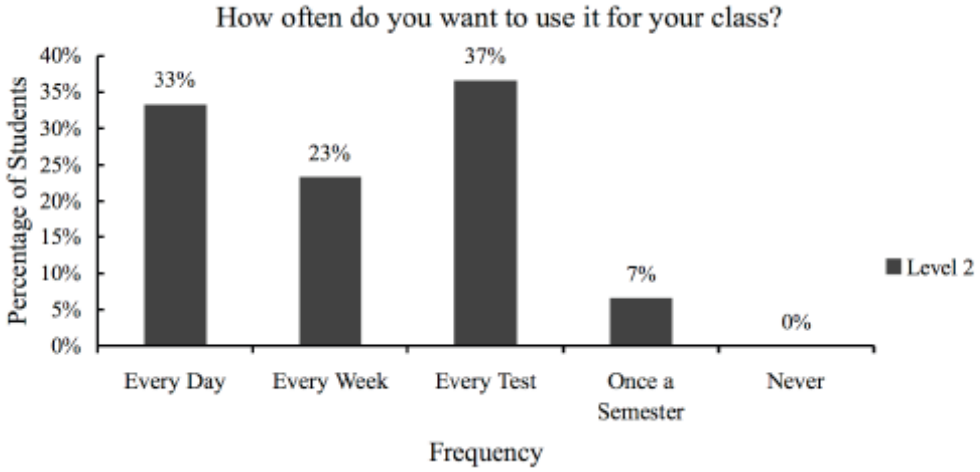

Figure 30. Student suggestions for frequency of use for Level 2.

Level 3. The Grammar 3 teachers varied in their opinions on implementation. Of the four teachers, two said it should be used as an oral component with each testing unit.

Two teachers wanted to use a variety of methods including self-assessment and feedback from the teacher. Another teacher wanted to only use the program for practice and not have any teacher feedback. Collectively, teacher recommendations included using the program every other week, every testing unit, or once a month. One teacher said it depends on the purpose. As for student feedback, 39 Grammar 3 students wanted to use the program in class for practice. 23 wanted to use it for out-of-class assignments, and 19 students thought it should be included on each grammar test (see Figure 31). As for the frequency of using *Grammatically Correct*, 24% of the Grammar 3 students wanted to use it every day, while 54% wanted to use it every week. 13% wanted to use it once every test, while 15% wanted to use the program only once a semester or not at all (see Figure 32).

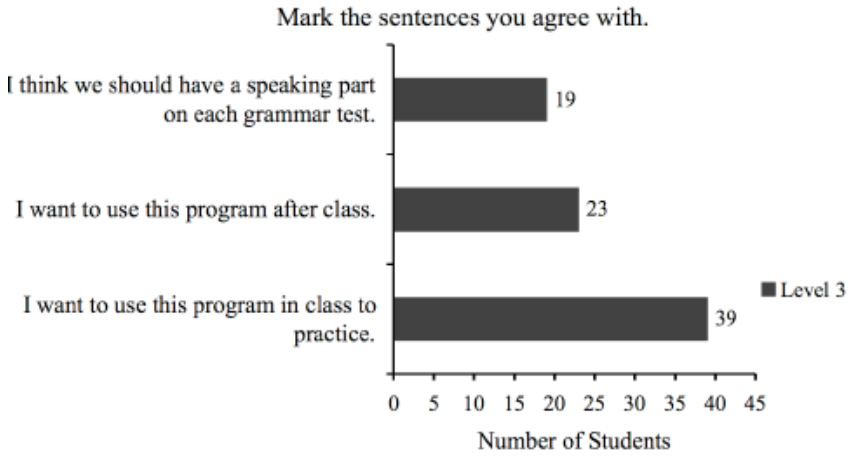

Figure 31. Student suggestions for methods of implementation for Level 3.

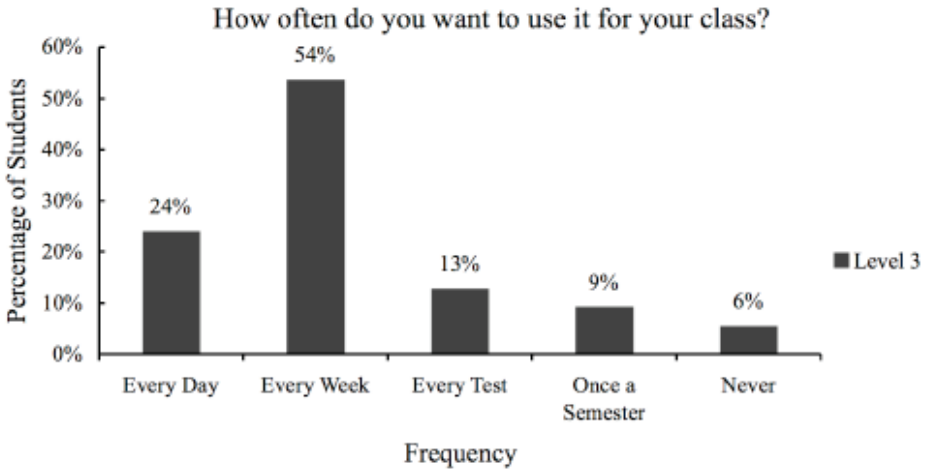

Figure 32. Student suggestions for frequency of use for Level 3.

Level 4. The Grammar 4 teacher said that both in-class sessions and out-of-class assignments, some graded and some not, would be effective for Level 4. He recommended using it every other week. Of the students in his classes, 15 wanted to use it for in-class practice, 10 wanted to use the program after class, and 11 thought it should be included on each grammar test (Figure 33). A large portion of Level 4 students want to use the program on a weekly basis, while 32% want to use the program only once during the semester (Figure 34).

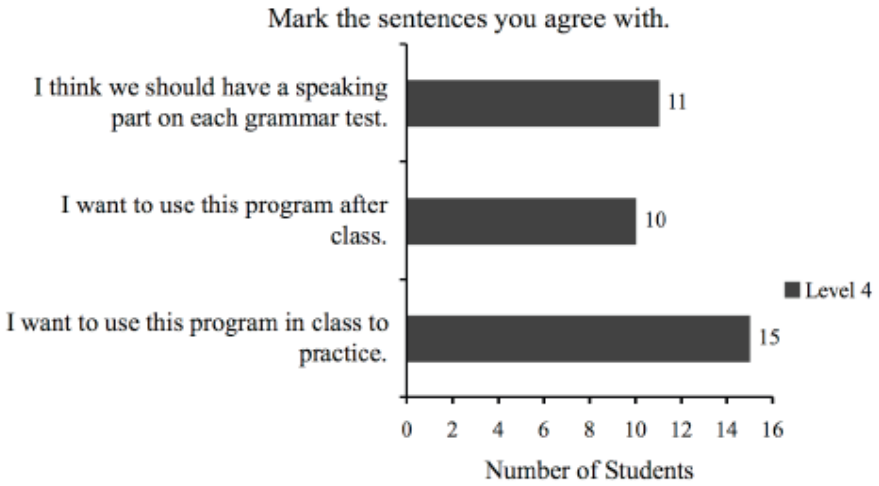

Figure 33. Student suggestions for methods of implementation for Level 4.

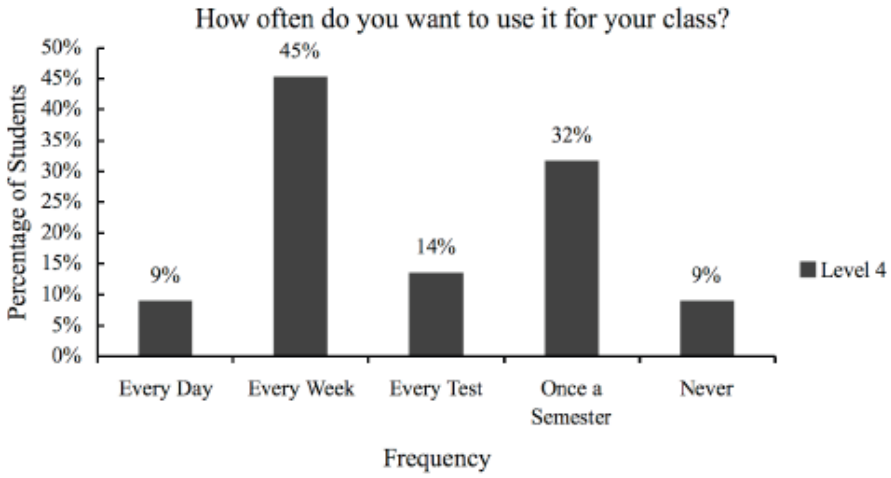

Figure 34. Student suggestions for frequency of use for Level 4.

The Need for Feedback. Looking at student responses across all levels, in-class practice was the most popular option. However, the majority of students (88%) also wanted their teacher to listen to them and give them feedback (see Figure 35).

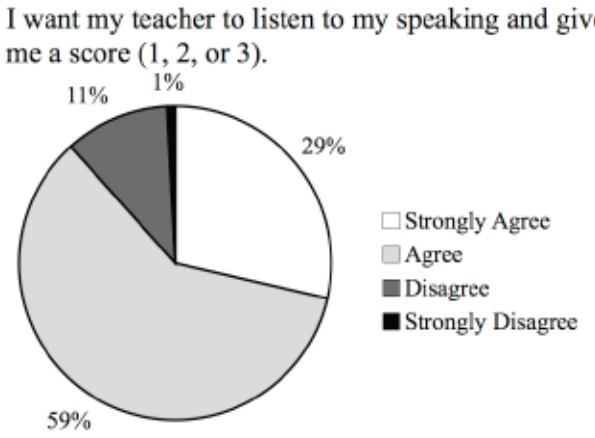

Figure 35. Student feedback on having their teacher to listen to their speaking.

One student wrote, "We need teacher or computer program to correct our mistake. If we only practice by ourselves without a T/A or teacher correct our grammar we won't learn." Another student wrote, "Give my score. But it is important to the teacher." The need for feedback is further supported by the classroom discussions. In the Level 4 class, eight

students said they wanted their teacher to listen to their responses, and there were similar comments in the Level 3 class discussion. Almost every student agreed that they wanted feedback, with about six students saying they preferred written feedback to simply a score. Most students also said they would rather get feedback from the teacher than peer- and self-assessment. In the Level 2 class, I asked them what they liked about *Grammatically Correct*, and six students said they liked that their teacher listened to them and gave them feedback. I received similar responses in the Level 1 class discussion, with six of the students in the Level 1 saying they liked to get feedback.

Summary of Recommendations for Implementation. Students and teachers were asked how often the program should be used. Among students of all levels, using the program on a weekly basis was the most popular choice for frequency of use. As for the teachers, they suggested using the program once every 2-4 weeks. Regarding the best way to use the program, it seems most teachers agree that the opportunity to practice is essential, and consider in-class practice sessions quite useful. Similarly, the most popular option for students was in-class practice, although a large group of the lower levels also wanted to use it as a testing component. Regardless of whether *Grammatically Correct* is used on actual tests, students alike expressed the need for teachers to listen to them and give feedback.

Question 3: What improvements should be made to the program?

Through this evaluation, it was also expected that suggestions for improvements to the program would naturally emerge. Several different topics were brought up students and teachers, including overall usability, specific changes to the website and recording application, training materials, and the issue of making custom task selections versus

having them already set up. The feedback from teachers regarding improvements is discussed first, followed by the feedback from students.

Feedback from Teachers. I designed the website to allow teachers to select their own tasks and custom-build selections. However, the teacher feedback shown in Figure 36 suggests that teachers would rather the selections be already set up for them.

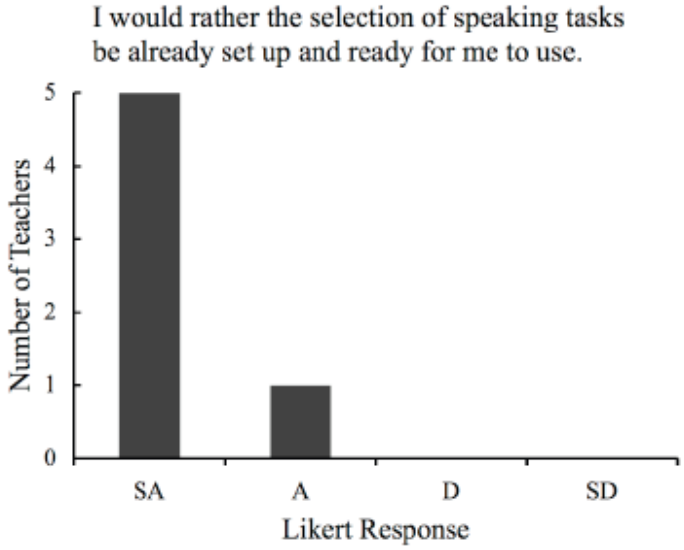

Figure 36. Teacher feedback regarding creating their own tasks and selections.

In teacher interviews, they said that having pre-made selections set up would be especially helpful to new teachers. One teacher said, "I think for me with my limited computer abilities and intuition, pre-made tasks are better. Again, if I felt absolutely confident is making my own task selections and setting it up, I could do it."

Another suggestion that came up during one of the interviews was to create better training materials. One teacher said that we need a training manual for teachers to use in the future, after the current ELC grammar coordinator and I have left. In the interview, we talked about possibly making a video that shows how to use the program, instead of only written instructions.

As for general usability issues, teacher feedback from the surveys and interviews included general computer problems, such as computers freezing and microphones not working. One teacher said that if she used the program more, perhaps she would feel more comfortable with it. Another said, "I think the program is well-designed. The computer lab could work on microphone issues."

Teachers also gave several suggestions on how to improve the website and recording application. The suggestions to the website mainly include changing the font-size and wording of some links and adding a few new links. As for the recording application, they said that it would be helpful to add a "Back" button for students to return to a previous task in the selection and re-record themselves without having to log in again, making practice sessions more flexible. One teacher also said that she would like to be able to fast-forward through the student's response to a different part of the audio clip. Currently, teachers have to listen from the beginning without being able to go directly to the middle or the end of a clip.

Feedback from Students. From the beginning, I have been concerned with usability for the students. On the student survey, 87% of students agreed that the program was easy to use (see Figure 37).

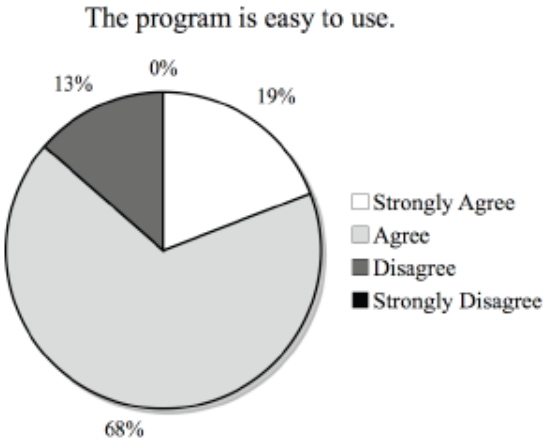

Figure 37. Student feedback on usability and clarity of directions.

As for the comments written on surveys, 13 were related to usability. Five students said the program was easy to use. One student wrote, "the program is helpful an easy to use," while another said, "That was easy to understand and the easy topics."

On the other hand, eight students on the survey said the program was difficult to use. "That program was really confused. I don't shure if the teacher explain well the instruction, but I my personal case, I didn't understand." Another example is, "The Instructions are not clear." Because of these eight survey comments, I tried to clarify what students found confusing about the program. However, in the class discussions I held with each level, the majority of the students said that the program was easy to use. I received no information about how to improve the program usability. This could have been because the students who did have problems were afraid to speak up. One suggestion made in the Level 4 class discussion was to be able to listen to the prompt, like with the Listening/Speaking tests.

Summary of Recommendations for Program Improvements. Most teachers preferred that task selections be pre-made and ready to use, rather than custom-designing

their own selections by choosing which tasks to use. As for usability, teachers suggested creating better training materials, as well as some minor changes on the website and recording application. The majority of students consider the computer application to be user-friendly, and they did not give any specific suggestions for improving usability.

Limitations of Evaluation

Like most evaluations, there are limitations due to resource constraints. The first limitation regards the first evaluation question which concerns the effectiveness of *Grammatically Correct* as an assessment instrument. As discussed in the "Purposes" section of the proposal, this evaluation is limited to measuring the teachers' and students' perceptions of using *Grammatically Correct* as a whole system and potential to use it for assessment in the future. The reason is that at this point, the program has not been used for actual assessment to a large enough extent. There are not enough student scores to analyze the effectiveness of actual assessment. Whether or not *Grammatically Correct* is actually effective in measuring grammatical ability is something that should be evaluated in the future, once the program is more fully implemented.

Another limitation concerns the reliability of the surveys used to collect data in this evaluation. Because of time and other resource constraints, no statistical reliability tests were run on the instruments. However, as described in the "Development and Validation of Instruments" section, other means were taken to validate and clarify wording and directions.

CHAPTER 6

CONCLUSIONS AND IMPLICATIONS

This final chapter of the project write-up discusses the overall conclusions gleaned from the evaluation, including the recommendations for changes that need to be made. The limitations of the project are presented next, followed by implications for further research.

Overall Conclusions

The purpose of this project was to develop an oral grammar assessment that is aligned with the ELC grammar course objectives, which could then provide specific diagnostic and achievement information to teachers and students. Since its initial design and development, the project has been implemented in a limited way in some of the ELC grammar classes. Because it was seen as an experimental program, teachers were encouraged, but not required, to try it in their classes as either practice or assessment. A formal evaluation, together with the informal evaluation conducted earlier, has provided helpful information regarding the perceptions of students and teachers, as well as suggestions for improvements.

Through the implementation and evaluation stages, teachers and students both expressed positive, encouraging feelings towards the assessment program. They liked the opportunity to use the grammar structures learned in actual speaking activities, which make the grammar seem less abstract. Furthermore, students felt it was beneficial to receive feedback on their oral production when teachers listened to them. Students were mostly concerned with the allotted preparation and speaking time, which can be easily remedied as tasks are written and revised.

As for implementation suggestions, the evaluation results show that both teachers and students want to use the program for practice, as well as assessment, on a regular basis. One major concern that arose from teachers, however, was the use of the program for actual assessment. While they reported on the survey that there is a potential for assessment, most teachers have been using it only as a practice tool, without scoring students' responses or giving very much feedback. Through interviews with teachers, it seems that many of them lack confidence because of their own unfamiliarity with using such a new program. Even among those are experienced users, some teachers lack confidence in their scoring judgments, which suggests that the effectiveness of the rubric is in question.

In order to reach the point where all grammar teachers are using the program and actual assessment is taking place, some changes could be made to help teachers become more confident with the program. First, having pre-made task selections to choose from, rather than requiring teachers to create their own, would reduce the time required to use the program and help teachers feel less overwhelmed with a new assessment tool. Another suggestion is to develop a short video demonstrating how to use the website and recording client. These two suggestions would be beneficial in providing more direction for teachers who are not as comfortable with learning new computer programs.

Project Limitations

As I have completed this project, an important limitation I see is the scope of language the program is able to elicit for assessment. Authentic language use often involves discourse as language users interact with each other in conversation. Because this assessment tool is designed to capture speaking samples on the computer, it is not

possible to create situations involving discourse, such as questions and answers, and we are limited to monologue speaking tasks. Furthermore, the speaking tasks are narrowly designed with a situation already provided, and students generally know which structure is being targeted. Therefore, the tasks are not necessarily effective at testing pragmatic competence.

The second limitation lies in the quality of the speaking tasks. Instead of a structured, solid grammar test, I created a flexible testing system that allows teachers or administrators to create their own tasks and selections and use the program as they please. We now have a task bank of approximately 430 speaking tasks, mostly written by me, but some by other teachers. Throughout the last few semesters, teachers have revised and added tasks at their leisure. Because of this, not all tasks are effective at eliciting the desired grammar structures.

Furthermore, some may see the program's flexible design as a weakness. Because I designed the program as a flexible system, I did not have a particular implementation plan in mind. Therefore, it has been difficult to know how best to use it in the ELC grammar classes, in terms of methods frequency and the type of assignments. However, the flexible design can also be seen a major strength, allowing ELC teachers and administration to set up assessments for each level, but at the same time giving flexibility to individual teachers who want it.

Finally, my technical resources were somewhat limited and I was not able to design the program exactly the way I wanted. I would have created the recording application with Flash, embedding it into the website. However, I did not have the

resources to learn Flash. Also, some of the suggestions for improving the program, such as being able to download student responses, are beyond my technical skills.

Directions for Further Research

As *Grammatically Correct* continues to be used as an practice tool and classroom assessment, further research should investigate the effectiveness of the rubric in measuring the mastery of specific grammar structures and the usefulness of its scores. When teachers are more confident with the rubric, they will probably be more confident in using the program for actual assessment. Individual tasks in the task bank should also be examined and refined to clarify the speaking prompts and topics, ensuring the elicitation of desired grammar.

This program also has the potential for research investigating grammar proficiency and testing. For instance, this assessment tool could be used to examine the relationship between oral production scores and written test scores. The program could also be used for research involving planning and speaking time. While there has been a group of studies investigating the effects of planning time on fluency, complexity, and accuracy in oral production, these studies only focus on overall grammar ability (Ellis, 2003; Ortega, 1999; Yuan & Ellis, 2003). Because this assessment tool focuses on individual grammar structures, it could be used to study the effect of planning and speaking time on accuracy of specific structures.

Another possibility extending beyond assessment is the potential for autonomous learning. Currently, the program is teacher-centered, requiring the teacher to set up selections to use according to the class schedule. However, it is possible to create practice selections that would be available to students for practice different grammar

structures at any time. In short, as the rubric and tasks are refined, this tool can be used for future research purposes, as well as for an effective practice and classroom assessment.

REFERENCES

Alderson, J.C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13,* 280-297.

Alderson, J.C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14,* 115-129.

Anderson, J., & Fincham, J. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology, 20,* 1322-1340.

Arkoudis, S., & O'Loughlin, K. (2004). Tensions between validity and outcomes: Teacher assessment of written work of recently arrived immigrant ESL students. *Language Testing*, *21*, 284-304.

Bachman, L.E., & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bailey, K. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*, 257-279.

Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing, 15,* 45-85.

Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teacher and Teacher Education, 15,* 253-271.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructor's classroom assessment practices: Purposes, methods, and procedures. *Language Testing, 21,* 360-389.

Cronbach, L. (1984). *Essentials of psychology testing* (4th ed.). New York: Harper & Row.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19,* 195-221.

DeKeyser, R. (1998). Beyond focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughy & M. Long (Eds.), *Handbook of second language learning* (pp. 313-348). Oxford: Blackwell.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL revitalization. *Studies in Second Language Acquisition, 13,* 431-469.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughy & M. Long (Eds.), *Handbook of second language learning* (pp. 313-348). Oxford: Blackwell.

Educational Testing Service. *iBT/Next Generation TOEFL Speaking Rubrics*. Retrieved January 27, 2007 from www.ets.org/Media/Tests/TOEFL/pdf/Speaking_ Rubrics.pdf.

Ellis, R. (1994). Introduction: Implicit and explicit language learning—An overview. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-31). San Diego, CA: Academic Press.

Ellis, R. (2001). Some thoughts on testing grammar: An SLA perspective. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, et al. (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 251-263). Cambridge: Cambridge University Press.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford: Oxford University Press.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*, 227-275.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly, 40*, 83-107.

English Language Center. (2006a). *Grammar binder.* Unpublished manuscript.

English Language Center. (2006b). *Speaking LAT scoring feedback sheets.* Unpublished manuscript.

Fuchs, M., & Bonner, M. (2006). *Focus on grammar 4* (3rd ed.). Longman: New York.

Fuchs, M., Bonner, M., & Westheimer, M. (2006). *Focus on grammar 3* (3rd ed.). Longman: New York.

Hymes, D. (1972). On communicative competence. In J. Gumperez & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 35-71). New York: Holt, Reinhart and Winston.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Englewood Cliffs, NJ: Prentice-Hall.

Kumaravadivelu, B. (2006). TESOL methods: Changing tracks, challenging trends. *TESOL Quarterly, 40*, 59-82

Larsen-Freeman, D., & Long, M.H. (1991). *An introduction to second language acquisition theory and research.* London: Longman.

Li, X.J. (1990). How powerful can a language test be? The MET in China. *Journal of Multilingual and Multicultural Development, 11,* 393-404.

Long, M.H. (1983). Native-speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4,* 126-141.

Long, M. H. (1988). Instructed interlanguage development. In L. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 115-141). New York: Harper and Row.

McNamara, T. (1996). *Measuring second language performance.* Essex, England: Addison Wesley Longman.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*, 5-11.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23,* 13-23.

Norris, L.M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*, 417-528.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21,* 109-148.

Ochs, E. (1979). Planned and unplanned discourse. In T. Givon (Ed.), *Syntax and semantics: Vol. 12. Discourse and syntax* (pp. 51-80). New York: Academic.

Purpura, J. (2004). *Assessing grammar.* Cambridge: Cambridge University Press.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18,* 27-67.

Rogers, T. (1991). Educational assessment in Canada: Evolution or extinction? *The Alberta Journal of Educational Research, 37*, 179-192.

Schoenberg, I. E. (2006). *Focus on grammar 2* (3rd ed.). Longman: New York.

Schoenberg, I. E., & Maurer, J. (2006). *Focus on grammar 1* (3rd ed.). Longman: New York.

Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback model for assessing foreign language learning. *The Modern Language Journal, 76,* 513-521.

Shohamy, E., Gordon, C., & Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct versus semi-direct oral tests. *The Modern Language Journal, 76,* 27-33.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Stufflebeam, D. (1995). The program evaluation standards. *ERIC Clearinghouse on Assessment and Evaluation Washington DC*. Retrieved November 16, 2006 from http://www.ericdigests.org/1996-1/the.htm.

VanPatten, B., & Oikennon, S. (1996). Explanation vs. structured input in processing instruction. *Studies in second language acquisition, 18,* 495-510.

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155). Cambridge: Cambridge University Press.

Weigle, S.C. (2002). *Assessing writing.* Cambridge: Cambridge University Press.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*, 1-27.

APPENDIX A: ELC GRAMMAR OBJECTIVES

Level 1 Objectives

Text:	*Focus on Grammar 1: Third Edition* **(**Yellow Cover)
Author:	Irene E. Schoenberg and Jay Maurer
Publisher:	Pearson Longman
Year:	2006
Workbook:	*Focus on Grammar 1: Third Edition* by Irene E. Schoenberg & Jay Maurer
Additional Materials:	Grammar 1 binder FOG 1: Teacher's Manual FOG 1: CDs FOG 1: Assessment Pack FOG 1: Transparencies More Grammar Practice Grammar in Context Grammar in Action

At the end of Level 1, the student should be able to use the following structures with **80 % accuracy**:

1. Imperative sentences (including negative)
2. "To be" present tense: statements, negative statements, yes/no questions and short answers, wh-questions, usage with "there"
3. "To be" past tense: statements, negative statements, yes/no questions and short answers, wh-questions, usage with "there"
4. Other verbs- present tense: statements, negative statements, yes/no questions and short answers, wh-questions
5. Other verbs- past tense (regular and irregular): statements, negative statements, yes/no questions and short answers, wh-questions
6. Present progressive tense: statements, negative statements, yes/no questions and short answers, wh-questions
7. Future tense with "be going to": statements, negative statements, yes/no questions and short answers, wh-questions
8. Singular & Plural Nouns
9. Possessive nouns
10. Articles: a/an, the, one/ones
11. Introductory Count and non-count nouns: some, any, much, many, a lot, a few, a little
12. How much/How many and quantity expressions
13. Noun & adjective modifiers
14. Comparative and superlative adjective forms (people and things)
15. Possessive adjectives
16. Adverbs of frequency: always, often, sometimes, usually, rarely, never

17. Subject and object pronouns
18. Prepositions of place: on, next to, from, at, in
19. Prepositions of time: in, on, at
20. Modals: can/can't
21. This/that and These/those

Level 2 Objectives

Text: *Focus on Grammar 2: Third Edition* (Blue Cover)
Author: Irene E. Schoenberg
Publisher: Pearson Longman
Year: 2006
Workbook: *Focus on Grammar 2 Workbook* by Samuela Edkstut
Additional Materials: Grammar 2 binder
FOG 2: Teacher's Manual
FOG 2: CDs
FOG 2: Assessment Pack
FOG 2: Transparencies

At the end of Level 2, the student should be able to use the following structures with **80 % accuracy**:

1. "To be" present tense: statements, negative statements, yes/no questions, wh-questions and short answers, usage with "there"
2. "To be" past tense: statements, negative statements, yes/no questions, wh-questions and shorts answers, usage with "there"
3. Other verb- present tense: statements, negative statements, yes/no questions, wh-questions and short answers
4. Other verb-past tense: statements, negative statements, yes/no questions, wh-questions and short answers
5. Present Progressive tense: statements, negative statements, yes/no questions, wh-questions and short answers
6. Future tense with will and be going to: statements, negative statements, yes/no questions, question-word questions and short answers
7. Non-action verbs
8. Imperative sentences
9. Possessive nouns
10. Proper Nouns
11. Count & Non-Count Nouns
12. Quantifiers: some, any, much, many, a lot of, a few, a little, too many, too few, too much, too many, too little, how much, how many
13. Articles: a/an, the, one/ones/it
14. Introductory Gerund and Infinitive usage
15. Descriptive Adjectives (place)
16. Comparative and Superlative adjective forms (people and things)

17. Possessive Adjectives
18. Adjectives with: enough, too, as + adjective + adjective, same/different
19. Adverbs of frequency: always, almost always, frequently, usually/often, sometimes, rarely/seldom, almost never, never and other time markers
20. Adverbs of Manner: -ly forms and other forms
21. Subject and object pronouns / direct and indirect objects
22. Possessive Pronouns
23. Prepositions of place: under, behind, on next to, between, near, in, in front of, in back of
24. Prepositions of time: in, on, at
25. Prepositions in addresses: on, at, on the, on the corner of
26. Modals of Ability and Possibility: can, could
27. Modals of Suggestions: Let's…, Why don't we…, Why don't you….?
28. Modals of Possibility: may, might
29. Modals of Permission: can, may
30. Modals of Requests, Desires, Offers: would like, would you like, would you please, I'd like
31. Modals of Advisability: should, ought to, had better
32. Modals of Necessity: have to, don't have to, must, mustn't
33. This/That and These/Those
34. Ordinal numbers

Level 3 Objectives

Text:	*Focus on Grammar 3: Third Edition* (Green Cover)
Author:	Marjorie Fuchs, Margaret Bonner and Miriam Westheimer
Publisher:	Pearson Longman
Year:	2006
Workbook:	*Focus on Grammar 3 Workbook* by Marjorie Fuchs
Additional Materials:	Grammar 3 binder FOG 3: Teacher's Manual FOG 3: CDs FOG 3: Assessment Pack FOG 3: Transparencies

At the end of Level 3, the student should be able to use the following structures with **80 % accuracy**:

1. "To be" verb and other verb- present tense: statements, negative statements, yes/no questions, question-word questions and short answers, non-action verbs
2. "To be" verb and other verb-past tense (regular & irregular): statements, negative statements, yes/no questions, questions word questions and short answers

3. Present Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
4. Past Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
5. Future tense: statements, negative statements, yes/no questions, question-word questions and short answers, future time clauses
6. Present Perfect tense: statements, negative statements, yes/no questions, question-word questions, and short answers, usage of "since," "for," "already," and "yet"
7. Present Perfect Progressive tense: statements, negative statements, yes/no questions, questions word questions, and short answers
8. Imperatives
9. Introductory Phrasal Verbs
10. Count & Non-count Noun review
11. Proper Noun review
12. Articles: Definite and Indefinite
13. Gerunds: as subjects and objects, after prepositions
14. Infinitives: after certain verbs, of purpose, with too and enough
15. Review of adjectives, adverbs of frequency, and adverbs of manner
16. Participial adjectives: -ing and –ed endings
17. Adjectives: Comparatives and Superlatives
18. Reflexive and Reciprocal Pronouns
19. Modals of Ability: can, could, be able to
20. Modals of Permission: can, could, may, do you mind if
21. Modals of Requests: can, could, will, would, would you mind
22. Modals of Advise: should, ought to, had better
23. Modals of Suggestions: let's, could, why don't, why not, how about
24. Modals of Preference: prefer, would prefer, would rather
25. Modals of Necessity: have (got) to, don't have to, must, must not , can't
26. Modals of Expectations: be supposed to
27. Modals of Future Possibility: may, might, could
28. Modals for Conclusions: must, have (got) to, may, might, could, can't

Level 4 Objectives

Text:	*Focus on Grammar 4*: *Third Edition* (Purple Cover)
Author:	Marjorie Fuchs and Margaret Bonner
Publisher:	Pearson Longman
Year:	2006
Workbook:	*Focus on Grammar Workbook* 4 by Marjorie Fuchs and Margaret Bonner
Additional Materials:	Grammar 4 binder FOG 4: Teacher's Manual FOG 4: CDs FOG 4: Assessment Pack

FOG 4: Transparencies

At the end of Level 4, the student should be able to use the following structures with **80 % accuracy**:

1. Present tense with "To be" verb and other verbs: statements, negative statements, yes/no questions, question-word questions and short answers
2. Past tense with "To be" verb and other verb (regular & irregular): statements, negative statements, yes/no questions, question-word questions and shorts answers
3. Present Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
4. Past Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
5. Present Perfect and Present Perfect Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
6. Past Perfect and Past Perfect Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
7. Future tense: statements, negative statements, yes/no questions, question-word questions and short answers
8. Future Progressive tense: statements, negative statements, yes/no questions, question-word questions and short answers
9. Negative and Tag questions
10. Responses with so, too, neither, not either, but
11. Additions with so, too, neither, not either, but
12. Imperative sentences
13. Phrasal verbs: Separable and Inseparable
14. Gerunds and Infinitives
15. Causatives: make, have, let, help, get
16. Adjective Clauses
17. Modals of Advisability and Obligation in the Past
18. Modals of Speculation and Conclusions about the Past
19. Passive Voice with Modals, Modal-like expressions and Causatives
20. Factual Conditionals in Present
21. Factual Conditionals in the Future
22. Unreal Conditionals in the Present
23. Unreal Conditionals in the Past
24. Direct and Indirect Speech (Noun Clauses)
25. Embedded Questions (Noun Clauses)
26. Present Perfect
27. Past Perfect
28. Past Perfect Progressive

APPENDIX B: DATABASE STRUCTURE

Classes Table: Stores information about each class.

Field name	Data type	Description
class_id	integer	Unique number that identifies a class
teacher_id	integer	The id number of the teacher, as set in the “teachers” table.
description	text	Description of the class for displaying on the web page
password	text	Class password for use by students

Images Table: Stores the prompt images that are uploaded.

Field name	Data type	Description
id	integer	Unique number to identify the image
image	binary object	The jpeg image itself
image_size	integer	The image's size in bytes

Publishing Table: Stores information about which class the assignment is for and what day/time it is to be available on the recording application.

Field name	Data type	Description
id	integer	Unique number to identify a published selection
class_id	integer	The id number of the class this selection is being published for
selection_id	integer	The id of the selection being published
start_time	date and time	Date and time students can begin
end_time	date and time	Date and time students can no longer access this selection
survey	integer	Set to one if the client is to present a survey to the students

Response Table: Stores the audio clips of the recorded student responses.

Field name	Data type	Description
id	integer	Unique identifier for the response
audio_clip	binary object	Flash audio of the student's response
student_id	integer	The id of the student as set in the “students” table
task_id	integer	The id of the task the student responded to
audio_clip_ size	integer	Size of the audio file

audio_clip_ mimetype	text	Mimetype of the audio file

Grammar Structures Table: Stores the list of available grammar structures.

Field name	Data type	Description
id	integer	Unique number identifying the struction
description	text	Brief description of the structure
family	text	Family of this structure

Student Table: Stores the names of each student, the task they completed, and the index of their audio clip.

Field name	Data type	Description
id	integer	Unique id of student record
class_id	integer	Class the student is participating in, as defined in the “classes” table
name	text	Student's name
time	date and time	Date and time the student did the selection
selection_id	integer	Id of the selection

Task Table: Stores information about each task.

Field name	Data type	Description
task_id	integer	Unique id for the task
level	integer	Level of the task
topic	text	Topic of the task
summary	text	Short summary of the task
preptime	integer	Time students will be allowed to prepare to spaek
speaktime	integer	Time allowed for students to speak
prompt	text	Task's prompt, to be displayed to students
structure	integer	Structure of the task, as defined in the “structure” table
prompt_ image_id	integer	Id of the image to display in the prompt, found in the “images” table
img_source	text	Image attribution information
teacher_id	integer	Id of the teacher who created this task

Task Selection Table: Stores information about each task that has been selected to a set of tasks, and which selection it belongs to.

Field name	Data type	Description
id	integer	
selection_id	integer	Selection id as defined in the "selections" table
question_ order	integer	Order the task will be displayed in the selection
task_id	integer	The task that will be displayed in the selection

Selections Table: Stores information about each selection, including the class that it is for, and if it a practice selection or a recorded quiz.

Field name	Data type	Description
selection_ id	integer	Unique identifier of the selection
teacher_id	integer	The id of the teacher who created the selection
description	text	Brief description of the selection, displayed to students in the client
graded	integer	Set to "1" if the selection is to be recorded to the server for grading

User Type Table: Stores the three types of users.

Field name	Data type	Description
id	integer	Id number of the user type
description	text	Description of the user type, for example "administrator," "teacher," or "demo"

User Table: Stores usernames and passwords.

Field name	Data type	Description
teacher_id	integer	Unique identifier for the teacher, used by records in other tables
teacher_ type	integer	Type of the teacher which indicates access level, such as "administrator"
first_name	text	
last_name	text	
username	text	Username that is used for teacher to log into the web site
password	text	
visible_	integer	Set to "1" if the students can see this teacher in the client

APPENDIX C: TEACHER SURVEY

Teacher Survey about Grammatically Correct

Dear Grammar teacher,

I really want to know your opinion about the *Grammatically Correct* computer program. Your feedback will be part of my formal evaluation which is necessary to complete my MA project. Please complete the following survey and return it to me by **December 1.** Thank you so much for your help and feedback.

Heather T

Consent to be a Research Subject

This research study is being conducted by Heather Torrie, a graduate student in Teaching English to Speakers of Other Languages at Brigham Young University. The purpose of this study is to evaluate the effectiveness of a newly developed oral grammar testing instrument.

You are asked to implement this testing instrument into your course syllabus. This will require you to take the extra time to create task selections and grade each response. This will take approximately 1 hour to create the selections, and up to 3 hours, over the course of the semester, to grade the responses. You are also asked to provide feedback through completing a questionnaire. The questionnaire will take approximately 20 minutes to complete. On the questionnaire you will be asked questions about your opinion of the test's usefulness and validity. Following the questionnaire, you will also be asked to participate in a focus group to talk about items on the questionnaire and any other concerns or comments you have regarding the online grammar testing application. The focus group will be scheduled at the convenience of all participating teachers. It will last approximately 30 minutes. All information will be kept confidential. After the study, all questionnaires and notes will be destroyed.

The risks in this study are minimal. There are also no direct benefits. However, your participation may help to improve this testing program. Participation in this research study is voluntary, and you have the right to withdraw from this study at any time.

If you have any questions about this research, contact Heather Torrie by phone at (801-422-4531 or by email at heather.torrie@gmail.com. If you have other questions about your rights as a research subject, contact Dr. Renea Beckstrand, the Chair of the Institutional Review Board, 422-3873, 422 SWKT, renea_beckstrand@byu.edu.

By signing and returning this questionnaire, you are giving your consent to participate in this study.

Signature: _____ Date: _____

Teacher Questionnaire about Grammatically Correct

Level you teach: _____

PART I: EFFECTIVENESS

1. ***Grammatically Correct* for Assessment Purposes.**

How effective is *Grammatically Correct* in helping students to practice the grammar structures they learn in speaking?

a. Very effective b. Somewhat effective c. Not very effective d. Not at all

How effective could *Grammatically Correct* be in actually assessing whether students have mastered the grammar structures they learn?

a. Very effective b. Somewhat effective c. Not very effective d. Not at all

2. **Benefits of *Grammatically Correct***

What do you like most about *Grammatically Correct*?

What is your biggest concern?

3. **The Speaking Tasks**

Please circle the answer which best describes your opinion of these statements. SA=Strongly Agree; A=Agree; D=Disagree; SD=Strongly Disagree

1. The tasks that I have used generally elicit the grammar structure in focus.	SA	A	D	SD	NA
2. The topics are appropriate for my students.	SA	A	D	SD	NA
3. The tasks in the task bank are well written.	SA	A	D	SD	NA
4. I would rather write my own tasks.	SA	A	D	SD	NA
5. I want to set up my own speaking assignments (Create a selection by choosing speaking tasks, set up the start/end times)	SA	A	D	SD	NA
6. I would rather the selection of speaking tasks be already set up and ready for me to use.	SA	A	D	SD	NA

4. **The Rubric:** Here is the suggested rubric for scoring the responses.

3	**Complete Mastery**: The structure is used correctly almost all of the time. There may be one or two mistakes in form, but overall the structure is used appropriately and with correct form most of the time. This shows that the structure has been mastered.
2	**Partial Mastery:** The structure is used correctly about half of the time. Sometimes it is used correctly, with the correct form. However, this accuracy is not consistent. Because the structure is used correctly only half of the time, it has been partially mastered.
1	**No Mastery:** The structure is used correctly almost never. There are many errors present in the form and way the structure is used. Because of these problems in accuracy, this structure has not been mastered.

In which ways did you use this rubric?

- ❑ I used the rubric to give my students feedback on their speaking assignments
- ❑ My students used the rubric to rate themselves and their classmates in class
- ❑ I used a different rubric
- ❑ I didn't use a rubric at all

Please circle the answer which best describes your opinion of these statements.
SA=Strongly Agree; A=Agree; D=Disagree; SD=Strongly Disagree

1. The rubric descriptions are clear and understandable.	SA	A	D	SD	NA
2. It is easy to distinguish between the three scores (1 – 2 – 3).	SA	A	D	SD	NA
3. The score I give my students helps me know if my students have met the objectives.	SA	A	D	SD	NA
4. I think the scores help students to know their level of achievement.	SA	A	D	SD	NA

What major concern(s) do you have about the rubric?

PART II: USE AND IMPLEMENTATION

5. **Current Use.** Think about your experience this semester with *Grammatically Correct.* Which of the following ways have you used the program in your grammar class? (Check all that apply)

   - ❑ An oral component along with each testing unit
   - ❑ Out-of-class assignments which are scored by the teacher and given back as feedback to students.
   - ❑ Out-of-class assignments that allow students to practice the grammar, but the teacher doesn't actually score their grammar.
   - ❑ In-class practice sessions in which students use self- and peer-assessment
   - ❑ In-class practice sessions for practice only without any kind of actual score
   - ❑ Print the tasks out and use them for in-class communicative practice activities
   - ❑ Other: _____

   How often did you use it? (How many times?)

6. **Implementation.** Given your experience, which of the following statements best represents your opinion on how *Grammatically Correct* **should be used for your level** in the future? (Check all that apply)

   - ❑ An oral component along with each testing unit
   - ❑ Out-of-class assignments which are scored by the teacher and given back as feedback to students.
   - ❑ Out-of-class assignments that allow students to practice the grammar, but the teacher doesn't actually score their grammar.
   - ❑ In-class practice sessions in which students use self- and peer-assessment
   - ❑ In-class practice sessions for practice only without any kind of actual score
   - ❑ Print the tasks out and use them for in-class communicative practice activities
   - ❑ Other: _____

   If you were to teach the same grammar course again next semester, which of these methods would you use? Please explain why.

For future semesters, **how often** would you recommend using *Grammatically Correct*?

a. Once every test unit	c. Every other week
b. Once a week	d. Other: _____

PART III: IMPROVEMENTS

What are the most important changes that should be made to the **website** (that teachers use to make selections and listen to students) to improve usability and convenience?

What are the most important changes that should be made to the **recording application** (that the students use) to improve usability and convenience?

PART IV: SUMMARY

Please write any other comments you have about the online oral grammar assessment tool that could help us make improvements.

APPENDIX D: STUDENT SURVEY

Thinking about Grammatically Correct

Consent to be a Research Subject

This research study is being conducted by Heather Torrie, a graduate student in Teaching English to Speakers of Other Languages at Brigham Young University. The purpose of this study is to evaluate the effectiveness of *Grammatically Correct*, the grammar speaking program, that you've been using this semester in your grammar class.

You are asked to complete a questionnaire. The questionnaire will ask you questions about what you think about the grammar tests. This will take about 10 minutes. Then, you might also be asked to participate in a group discussion. This means that the researcher will come to your class and talk to you and your classmates for about 15 minutes. Here, you can talk about the things on the questionnaire and about the grammar speaking tasks. All information will be kept confidential. Your name will not be on any information. After the study, all questionnaires and notes will be destroyed.

The risks in this study are minimal. There are also no direct benefits. However, your answers and opinion can help to improve this grammar program. Participation in this research study is voluntary, and you have the right to withdraw from this study at any time. This means that you don't have to answer these questions if you don't want to.

If you have any questions about this research, contact Heather Torrie by phone at 801-422-4531 or by email at heather.torrie@gmail.com. If you have other questions about your rights as a research subject, contact Dr. Renea Beckstrand, the Chair of the Institutional Review Board, 422-3873, 422 SWKT, renea_beckstrand@byu.edu.

By signing and returning this questionnaire, you are giving your consent to participate in this study.

Signature: _____ Date: _____

Background Information:

Level _____
Teacher _____
Country: _____
Native Language: _____
Age: _____

Think about the grammar speaking program that you used in your grammar class. Read the sentences below. Circle the word (Strongly Agree, Agree, Disagree, or Strongly Disagree) *that tells <u>how you feel about the sentence.</u>*

1. I like using this program.

Strongly Agree Agree Disagree Strongly Disagree

2. The program is easy to use.

Strongly Agree Agree Disagree Strongly Disagree

3. The directions are clear (not confusing).

Strongly Agree Agree Disagree Strongly Disagree

4. I have enough time to plan.

Strongly Agree Agree Disagree Strongly Disagree

5. I have enough time to speak.

Strongly Agree Agree Disagree Strongly Disagree

6. I like the speaking topics.

Strongly Agree Agree Disagree Strongly Disagree

7. The program helps me practice the grammar I learn in my grammar class.

Strongly Agree Agree Disagree Strongly Disagree

8. It helps me to listen to my grammar.

Strongly Agree Agree Disagree Strongly Disagree

9. I can hear when I make a mistake.

Strongly Agree Agree Disagree Strongly Disagree

10. I want my teacher to listen to my speaking and give me a score (1, 2, or 3).

Strongly Agree Agree Disagree Strongly Disagree

1. How often do you want to use it for your ELC grammar class? *Circle one.*
   a. Every day
   b. Every week
   c. For every test
   d. One time each semester
   e. Never

2. Mark the sentences you agree with: *(You can mark more than one answer)*
   - ❑ I want my teacher to listen to my recording and give me a score.
   - ❑ I want to use this program **in class** to practice.
   - ❑ I want to use this program to practice **after class**.
   - ❑ I think we should have a speaking part on each grammar test.

3. What did you most like about the *Grammatically Correct* (the grammar speaking program)?

4. What did you not like about the *Grammatically Correct* (the grammar speaking program)?