Computer-Mediated Grammar Practice and its Effect on Different Language Tasks

FRANZISKA lys
Northwestern University

ABSTRACT
When and to what extent one should teach grammar to language learners has been discussed extensively. Even though grammar continues to hold an important place in language teaching, there is still much controversy with regard to what works best. This article looks at the question of computer-mediated grammar teaching and its effect on language acquisition over time. In particular, it reports the results of a year-long empirical study of the effectiveness of extended online practice of L2 grammar on the overall development of the L2 in learners of German. The results of the study suggest that prolonged online grammar practice translated into higher scores on chapter exams. However, no effect was found with regard to written work or oral chats and we concluded that prolonged online grammar practice did not have an effect on the quality of open-ended tasks such as these over the course of a year. Implications for language learning are discussed, exploring possible explanations on the short-term effect of online grammar instruction and the apparent lack of cumulative effect on more open-ended tasks.

INTRODUCTION
This study adds to the existing discussion on the explicit teaching of grammar by investigating computer-mediated grammar practice and its effect on language acquisition over time. In particular, it reports the results of a yearlong empirical study exploring the effectiveness of extended online L2 grammar practice on the overall development of L2 in learners of German at the intermediate level of language instruction.¹

When and to what extent one should teach grammar to L2 learners has been discussed extensively in theoretical and pedagogical literature (see for example Nassaji & Fotos, 2004). Questions such as “Should grammar be taught?”, “How and when?”, “Are grammar drills useful?”, and “How do we integrate grammar teaching into a communicative approach?” are quite common. Although grammar continues to hold an important place in language teaching, there is consider-
able controversy with regard to what works best (e.g., Ellis, 2006). The belief that students will develop a solid grammar foundation simply by being exposed to enough structured input is often met by the criticism that students who were taught through a communicative, input-based approach are not acquiring enough grammar knowledge. Giving more explicit grammar lessons, however, does not seem to be the solution either: “It may be a battle trying to get students to assimilate grammar rules entirely through input/communicative activities, but many instructors are not convinced that students should receive more explicit instruction when it comes to studying certain, more complicated grammar points” (Katz & Blyth, 2008, p. 3). The reason for this, as Katz and Watzinger-Tharp (2008) point out, are the various conceptualizations of grammar and the resulting teaching practices, which make it difficult for researchers as well as language instructors “to reach an agreement on how best to promote the acquisition of grammar” (p. xv). Ellis (2006) presents a series of personal interpretations of what the research to date has shown, but cautions that many of these statements are open to challenge. He points to the need for more research that “addresses to what extent and in what ways grammar instruction results in implicit knowledge” (p. 103).

LITERATURE REVIEW

Explicit and Implicit Teaching and Learning

The effectiveness of explicit and implicit teaching and learning in second language acquisition is an important issue that has been extensively discussed in the literature (e.g., DeKeyser, 2003; Doughty & Long, 2003; Norris & Ortega, 2000). At the center is the question of whether an adult learner can acquire a language fully with the same implicit learning style as a child acquiring a first language. To answer this question it is important to understand the difference between explicit and implicit teaching and learning and the role each play in second language acquisition. Explicit teaching and learning refers to the conscious learning and application of rules. This may include situations where rules are given or explained to the learner, where attention is drawn to the rule, or where learners are encouraged to discover underlying rules in the input themselves (Doughty & Long, 2003). Explicit instruction generally involves a qualified instructor who provides a variety of input that contains the rules of the target grammar and also ample practice opportunities with feedback (written or oral). In implicit teaching and learning no rules are taught. New structures in the target language are acquired by the learner through exposure to frequent input and output opportunities without a conscious effort to discover a rule. Spontaneous comprehension or production is believed to be the result of implicit language knowledge.

Comparing the research on explicit and implicit learning, DeKeyser (2003) points out that implicit learning of abstract rules may be very difficult to test (i.e., how much use did the learner make of rule knowledge even if such knowledge was not explicitly taught) and he concludes that we should be “very skeptical about the possibility of implicit learning of abstract nature, at least by adults” (p.
Furthermore, it may be difficult to obtain pure measures of either explicit or implicit learning, as these concepts may not represent clear and distinct categories. In a recent volume aimed at providing a broad view on explicit/implicit learning as modifiers in second language acquisition, Sanz (2011) sees the distinction as a continuum rather than a contrast “to distinguish between processes under attention, those that require awareness and those that take place in the absence of both” (p. 6).

Studies dealing with language learning as a result of some type of explicit teaching and learning have also focused on the role of related concepts such as noticing, awareness, attention, and focus on form. Many of these studies report a positive effect for explicit instruction either through the teaching of grammar rules, error corrections, or some type of input enhancement (DeKeyser, 2003). However, the precise role of instruction in developing language proficiency and whether it can affect implicit language knowledge is still under investigation (see for example Ellis, 2005). There is some evidence which suggests that a number of grammatical forms may be resistant to external manipulation by teachers and by learners (Macaro & Masterman, 2006; VanPatten, 2011).

For the most part, methods of foreign language instruction have moved away from explicitly focused grammar teaching to methods that focus on communicative competence and task-based learning. However, these two approaches are not incompatible and many instructors find that some type of grammar teaching is still beneficial in learning a language. Lightbown, Spada, and White (1993) for example point out that learners from immersion programs where no grammar was taught had significant shortcomings with regard to accuracy. Therefore, combining the two approaches may be a useful solution to becoming a well-rounded speaker of a foreign language. This may be especially true in the case of adult learners, who, for the most part, have the cognitive maturity to attend to rules and various forms.

The question of whether explicit grammar learning is beneficial for German as L2 is controversial as well. As summarized in Eckerth and Tschirner (2010), some research has shown that a focus on form can speed up the learning of German (e.g., Roehr, 2007; Schulz, 2002). Other research suggests that explicit instruction yields little results, as there might be certain developmental sequences in the acquisition of German as L2 such as word order, tense, and the case system, that follow a more defined developmental path regardless of instruction received (Meerholz-Härle & Tschirner, 2001; Diehl, Pistorius, & Dietl, 2002).

**Computer-Mediated Grammar Practice**

Using computers to practice and reinforce grammar is not new. Some of the earliest computer-assisted language learning environments focused solely on supplemental grammar where short answers or multiple-choice exercises gave learners a chance to practice correct grammatical forms. With the advances of technology, the approach to teaching and learning grammar has changed considerably from blended learning environments to fully online programs with more sophisticated tasks and feedback. In these programs, grammar instruction has become more
aligned with the goals of the communicative-focused language learning approach, realizing that “grammar exercises need to require more than single word or phrase answers” (Godwin-Jones, 2009, p. 5). While still encouraging a focus on form, many of the CALL activities today are fully grounded in task-based learning activities where the transmission of meaning through interactions with other speakers of the language is the focus (Godwin-Jones, 2009).

The body of research concerned with the format and effectiveness of online grammar teaching and learning is quite large, beginning with computer-driven tutorials and drills for teaching grammar to adults, to more collaborative and explorative learning environments (e.g., Bloch, 2009; Kessler, 2009; Sauro, 2009). Many studies suggest that computers can be beneficial for learning grammar as they can provide additional practice in a variety of contexts. Nagata (1996), for example, compared the effectiveness of Nihongo-CALI (which uses natural language processing to provide a more sophisticated error analysis) to written feedback in workbook instructions. The results showed that intelligent computer feedback was more effective than simple workbook answer keys for developing learners’ grammar. Nutta (1998) examined whether there was a difference in the acquisition of specific grammar points for students taught in a teacher-directed class versus those taught with computer-based instruction. She concluded that computer-based instruction is an effective method of teaching L2 grammar, especially when using grammar in open-ended tasks. More recently, Reima (2005) investigated the use of online grammar practice outside of the classroom as a supplement to in-class teaching practices in an EFL classroom. The results showed that online grammar practice improved students’ learning and mastery of English grammar. And finally, Faizah (2009) found that students using an internet-based grammar performed better on certain grammatical items and made fewer errors in their essays than those using conventional grammar instruction.

A BLENDED LEARNING ENVIRONMENT FOR INTERMEDIATE STUDENTS OF GERMAN

The second-year language program at Northwestern University uses an integrated blended learning approach where face-to-face classroom teaching is combined with computer-mediated activities. Reading, writing, speaking, and listening comprehension are emphasized throughout the year. Because of the diverse student background (about forty-five percent of the students in the program come from our first-year program; about fifty-five percent are entering freshmen from various high school programs around the country), we use an online review grammar called Intermatik to give all students a chance to review basic grammar points before being exposed to more advanced grammatical structures. The online grammar practice is fully integrated with the cultural multimedia program and the cultural and literary reading material used throughout the year.
**Intermatik, an Online German Grammar**

*Intermatik* is an online learning environment with a series of online grammar lessons modeled after key concepts developed in Fischer and Farris (1995, 1999) and Fischer (1997). The instructional design of the online learning environment was conceived by borrowing from the pedagogical concept developed for *Libra* (Fischer & Farris, 1995). Similar to a *Libra* lesson, *Intermatik* was designed to a) present information through text, audio, graphics, and digital video, and b) practice and verify application and comprehension through the use of several kinds of question formats and free writing tasks.

Working with *Libra* had shown us how important it was to begin the development process by “addressing questions of instructional design and then implement appropriate pedagogical features in support of that design” (Fischer & Farris, 1999, pp. 77-78). The goal of *Intermatik* was to create a learning environment to be used predominantly outside of class to study grammar points, listen to example sentences, and practice the new grammar forms in an interactive format focusing on form. This would free up class time during which students would be engaged in interactive task-based activities using the new grammar points through a variety of input-based and output-based exercises emphasizing grammatical accuracy, meaning, and conversational fluency (Ellis, 2002).

*Intermatik* is composed of twelve chapters each subdivided into five distinct parts: an introduction to a culture topic (*Kultur*), a vocabulary section (*Vokabular*), a grammar section (*Grammatik*), a review section (*Wiederholung*), and a resource section (*Ressourcen*).

The culture section introduces the students to the new theme and improves their cultural knowledge about the new theme. It includes a set of personal questions to engage students in the topic, a set of questions pertaining to cultural information which students research on the Internet, a broader (sometimes controversial) discussion question based on the topic, and a small set of additional cultural tasks for small group work or class discussion. These task-based exercises were drawn from a variety of authentic sources.

The vocabulary section contains a topical vocabulary list and a set of exercises to practice the new words and expressions. The first vocabulary exercise is usually a drag-and-drop exercise to help students recognize the English and German equivalents. The remaining exercises are designed to practice and solidify the new expressions in context. Many of the vocabulary exercises are questions with partial words missing. After filling in the new words, students are asked to bring these questions to class for partner work.

Each grammar section presents grammar explanations and several practice sessions that could also be used as test sessions. The concise, easy-to-follow grammar explanations are accompanied by example sentences spoken by native speakers. Each grammar explanation, in turn, is followed by a variety of exercises with basic feedback to offer the learner opportunities to practice each new point. Exercise formats are multiple-choice questions, check-list questions, pull-down menus, open-ended responses ranging from short answers to full sentence
rewrites, and various drag-and-drop options. All examples and exercises are based on the cultural theme in each unit.

Each review section contains six different review exercises designed to review/test the vocabulary and new grammar in context. The first four exercises follow a similar format as the exercises in the grammar part. The fifth exercise is always a global listening comprehension task with accompanying questions and the sixth exercise contains a video clip to elicit open-ended responses for a more creative use of the language integrating all of the grammar points practiced.

Learners also have a resource section at their disposal to review active vocabulary lists, irregular verbs, grammar tables, as well as read site explanations and technical help. (See Appendix A for an overview of each section in Intermatik.)

When a student logs into Intermatik, the program tracks the work she/he performs. It records how long each learner is engaged in practicing grammatical forms, what progress the learner makes with each exercise, and how well the learner understands the material in the end. The tracking results are available to the learner and the instructor and are part of the feedback on learning.

In designing the various exercises and tasks in Intermatik, we paid special attention to the type of task in each part moving from controlled, to semi-controlled, to open-ended activities. The culture section and the review section contain mostly open-ended activities to encourage language use on a variety of levels. Many of the culture tasks students prepare online are designed for follow-up partner work in class. Since the Intermatik program does not offer practice in pronunciation, the follow-up work done in class gives each instructor a chance to selectively practice pronunciation with each student as needed. In addition, the various tasks in the online learning environment were designed to highlight top-down and bottom-up processing strategies.

THE STUDY

The purpose of this study was to evaluate student interactions with the online grammar and its effect on the overall development of L2. In the study, I address the following research questions:

1. **Overall improvement of L2**: Does the regular practice of grammar lessons online improve the learning of German over the course of a year?
2. **Improvement across different task types**: How effective is the online grammar practice with regard to different tasks? What language skill improved the most at the end of the practice period?
3. **Interaction with the online grammar program**: How do students interact with the online grammar program? How often do they engage with the online grammar and how much time do they spend with each exercise?
4. **Self-assessment**: Did students feel that the use of the online grammar helped them learn? What was most beneficial in their view?

For the study we collected two sets of data. The first data set assessed overall language abilities (spoken, written, and a c-test). Tests were administered in the fall and then again in the spring. For the second data set, we tracked students’ work in the online grammar (completion rate and the time spent using the online
resources) during fall and spring quarter. We then compared data set one to data set two to ascertain the short-term effect of grammar instruction and its cumulative effect.

**Methodology**

**Participants**

The subjects for the study were students enrolled in the second-year German language program at Northwestern University. There were seventy-four students enrolled in the fall quarter, sixty-eight students in the winter quarter, and sixty-five students in the spring quarter. Of those, fifty-three students completed all three quarters and were included in the study. The final group of fifty-three students was comprised of twenty-seven males and twenty-six females. Fifty were undergraduate students (eleven freshmen, sixteen sophomores, fifteen juniors, eight seniors) and three graduate students. The students came from a variety of schools: thirty-eight were enrolled in the Weinberg College of Arts and Sciences, seven in the Bienen School of Music, three in the School of Communication, two in the McCormick School of Engineering, and three in the Kellogg School of Management. Twenty-four students had taken our first-year German program and twenty-nine students had been placed in second-year German by our online placement exam.

**Data Collection**

The study was a one-group pretest-posttest design with assessment periods during the fall and spring quarter (first and third quarter of second-year German). We administered the following set of assessments in fall and spring: a modified oral proficiency test (conversational interactions among peers), a written composition, and a modified cloze-test called a c-test. In both quarters we also collected survey data and class performance data. In addition, we tracked student time online and recorded the performance on each grammar exercise. These assessment data allowed us to conduct a set of analyses using matched-pairs t-tests and regression analysis examining learning practices, student attitudes, and performance measures.

**Overall Language Proficiency Assessments**

**Format of C-test.** To assess grammar comprehension and overall language proficiency, we used a c-test consisting of three short paragraphs of increasing difficulty on a variety of topics (selected from authentic texts) with a total of seventy-five blanks for students to complete. In each paragraph, the first and the last sentences were complete. For the sentences in the middle, the c-principle was applied: the second half of every second word was omitted (see Appendix B). For each blank, students were asked to complete the word so that the passage would make sense in German. Only entirely correct restorations were counted (Klein-Braley, 1997). We used the c-test principle for a variety of reasons: a c-test is easy to administer, students have few problems with the format, and grading is straightforward.
Since its first use in the early 1980s (Raatz & Klein-Braley, 1981) the c-test principle has been applied in numerous contexts across many different languages. Although some researchers have questioned what the test really measures (e.g., Eckes & Grotjahn, 2006), many agree that it provides a good assessment of general language ability. Of course, it is critical that texts be carefully selected and pre-tested for their validity (Jafarpur, 2002).

In establishing the construct validity of our c-test, we randomly distributed the three paragraphs to 144 students in classes across three years of language instruction the previous spring. What we found was a regular progression from low to high scores across increasing proficiency levels. The average score for the first-year students was 24 (n=59); the average score for the second-year students was 33 (n=69); and the average score for the third-year students was 50 (n=16).

**Oral proficiency test.** To assess the oral proficiency of the subjects we decided not to administer a regular oral proficiency test. The oral test design we used instead consisted of a chat or conversation among three or four students who were given a conversation topic in English (see Appendix C). They had fifteen to twenty minutes during which time they were asked to discuss the topic with the other students in German. Two instructors were present during the chat and they graded each subject independently on a scale of 0 (a clear demonstration of incompetence) to 3 (demonstrates strong command of the language) according to pronunciation and fluency, vocabulary, communicative effectiveness, and grammar. All oral performances were video taped. If the final two scores for a subject were more than one point apart, the grading was resolved by one additional grader.

This form of oral assessment was different from an interview as the responsibility for the flow of the conversation was handed over to the students entirely. There was no instructor asking leading questions or structuring the interview. This type of task mimicked the group conversations we had practiced in class on several occasions. During the test, we allowed for a five-minute warm-up period of speaking before assessing each student. We also realized that some subjects might struggle not only with the linguistic forms but also with the discursive nature of the task: it might be difficult for them to “break in” and be heard because they were shy or were never taught the linguistic tools. Pre-chat instructions included sample expressions to be used during the chat to make sure all subjects had a chance to speak.

**Written composition.** The written competence of each subject was assessed through an in-class writing task. Subjects were asked to write a letter to their parents (for the specific topic see Appendix D). The allotted time was twenty-five minutes and the requested essay length was ninety to one hundred words (or roughly 15 to 16 sentences). We did not allow the students to use a dictionary or ask instructors for help.

The final essays were assessed globally in three categories: content and effectiveness of the message, vocabulary and idiomatic phrasing, and grammar accuracy. We used a scale from 0 to 3 adapted from the evaluation measures used for grading AP exams (0=suggests incompetence, 1=suggests competence, 2=dem-
onstrates competence, and 3=demonstrates excellence). Each composition was graded independently by two instructors in the program. We used a double-blind process. If the final two scores of a composition were more than one point apart, the grading was resolved by one additional grader.

**Grammar Homework Assessment**

Part of the student homework for the class consisted of daily work online in *Intermatik* learning about German culture, studying new vocabulary and grammar explanations with example sentences, and practicing the material through a variety of interactive exercises. We tracked student interactions with the online material during fall and spring quarter through the program’s built-in tracking program as seen in Figure 1.

**Figure 1. Tracking record of online grammar work**

<table>
<thead>
<tr>
<th>Student Number</th>
<th>Chapter and Section Number</th>
<th>Exercise Number</th>
<th>Percentage Correct</th>
<th>Date and Time of Submission of Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 1</td>
<td>Exercise 2</td>
<td>33%</td>
<td>04/07/02_07:51:57 PM</td>
</tr>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 1</td>
<td>Exercise 2</td>
<td>50%</td>
<td>04/07/02_07:54:25 PM</td>
</tr>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 1</td>
<td>Exercise 2</td>
<td>50%</td>
<td>04/07/02_07:54:32 PM</td>
</tr>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 2</td>
<td>Exercise 1</td>
<td>16%</td>
<td>04/07/02_08:04:55 PM</td>
</tr>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 2</td>
<td>Exercise 1</td>
<td>66%</td>
<td>04/07/02_08:06:51 PM</td>
</tr>
<tr>
<td>snk211</td>
<td>Kapitel 12 Section 2</td>
<td>Exercise 1</td>
<td>100%</td>
<td>04/07/02_08:07:54 PM</td>
</tr>
</tbody>
</table>

The program kept track of the initial percentage of correct answers, whether the student continued to work on the exercise, the final percentage of correct answers, as well as date and time an exercise was submitted.

**Surveys**

We also asked students to fill out a survey called Approaches to Study Inventory (ASI). This survey, adapted from Entwistle (1988), was designed to categorize each student’s learning preference. In addition, students filled out several shorter surveys designed to collect information on students’ backgrounds and students’ attitudes towards learning German and the online grammar material.

**RESULTS**

**Oral Proficiency**

The oral proficiency of forty-six subjects was assessed in the fall and spring quarter through recorded oral group chats. The average grade for those students at the end of fall quarter was 1.646 and at the end of spring quarter was 2.108. The average overall gain was 0.462 points over the pre- and post-test which is significant
at conventional levels \((n=46, p=0.000)\) using a paired two-sided t-test. A lower score was given to 24% of the students. In Table 1, I present the average results for the oral chat for fall and spring as well as the change between fall and spring (Mean, SD, and p-value).

Table 1. Oral assessment results: Fall and spring

<table>
<thead>
<tr>
<th></th>
<th>Fall</th>
<th>Spring</th>
<th>Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Mean</td>
<td>1.646</td>
<td>2.108</td>
<td>0.462</td>
</tr>
<tr>
<td>SD</td>
<td>0.582</td>
<td>0.437</td>
<td>0.551</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

Written Assessment

Forty-nine students completed the composition in the fall and in the spring quarter. The average composition grade at the end of fall quarter was 1.708 and at the end of spring quarter was 1.867. The average overall gain over the pre- and post-test was 0.159, which is significant at conventional levels \((n=49, p=0.037)\) using a paired two-sided t-test. Thirty-three percent (sixteen out of forty-nine students) showed a negative gain from fall to spring quarter. In Table 2, I present the average composition results for fall and spring as well as the change between fall and spring (Mean, SD, and p-value).

Table 2. Written assessment results: Fall and spring

<table>
<thead>
<tr>
<th></th>
<th>Fall</th>
<th>Spring</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Mean</td>
<td>1.708</td>
<td>1.867</td>
<td>0.159</td>
</tr>
<tr>
<td>SD</td>
<td>0.567</td>
<td>0.509</td>
<td>0.519</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;0.040</td>
<td></td>
</tr>
</tbody>
</table>

C-Test

Forty-six students took the c-test at the end of fall quarter and at the end of spring quarter. Students in the fall restored on average 27.283 blanks (out of 75) correctly across the three paragraphs (P1, P2, P3). In the spring, they restored on average 40.391 blanks correctly. This represents an average gain of 13.109 blanks over the pre- and post-test which is significant at conventional levels \((n=46, p=0.000)\) using a paired two-sided t-test. These results were higher than the results we observed in the validity test we administered the year before. The average gain from first to second year in the validity test was nine points. There were no negative gains in the c-test. Table 3 represents the average c-test results for fall and spring as well as the change between fall and spring (Mean, SD, and p-value).
Table 3. C-test results: Fall and spring

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Fall scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>9.239</td>
<td>10.500</td>
<td>7.543</td>
<td>27.283</td>
</tr>
<tr>
<td>SD</td>
<td>3.591</td>
<td>3.758</td>
<td>2.510</td>
<td>8.074</td>
</tr>
<tr>
<td>Spring scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>13.652</td>
<td>15.522</td>
<td>11.217</td>
<td>40.391</td>
</tr>
<tr>
<td>SD</td>
<td>3.288</td>
<td>2.787</td>
<td>3.747</td>
<td>8.560</td>
</tr>
<tr>
<td>Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.413</td>
<td>5.022</td>
<td>3.674</td>
<td>13.109</td>
</tr>
<tr>
<td>SD</td>
<td>3.531</td>
<td>3.370</td>
<td>3.300</td>
<td>7.430</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Use of Online Grammar**

When analyzing the interaction of each student with the online grammar, we found two significant patterns: percentage of correct answers at quitting point and time at which students used the online grammar. Forty-three students quit the work only after they had achieved a relatively high score, namely 80% or more (0-79%=n10; 80-89%=n13; 90-100%=n30). With regard to time, our analysis yielded two significant patterns: students who primarily used the online grammar during the A.M. period and those who used the online grammar during the P.M. period. Students who primarily worked with *Intermatik* during the P.M. hours consistently outperformed the A.M. students as seen in Figure 2.
Regression Results

We did not find any significant difference between usage pattern of the online grammar and the oral chat performance, the written composition, or the c-test. However, we found a highly significant correlation between usage pattern of the online grammar work and the chapter quizzes as summarized in Table 4.

Table 4. Correlation results for usage pattern of the online grammar and quiz grades

<table>
<thead>
<tr>
<th>Quiz # 1</th>
<th>Quiz # 2</th>
<th>Quiz # 3</th>
<th>Quiz # 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>.361 (p=.016)</td>
<td>.381 (p=.005)</td>
<td>.369 (p=.007)</td>
<td>.310 (p=.019)</td>
</tr>
</tbody>
</table>

(P values less than .05 are statistically significant.)

Survey Results

Tabulating the results of the survey on use and benefit of the online grammar showed that students generally felt positive about *Intermatik* as a learning tool. About 24% of the students logged on to the online material every day, while fifty percent of the students logged on every other day. Judging by a 4-point Likert scale, the students found the interactive exercises and the online resources, such as grammar tables and vocabulary help, the most important (3 and 3.5). Feedback to the exercises was not as important as having good grammar explanations and a variety of practice and review exercises. Students indicated that they felt the online grammar helped improve grammar skills, followed by writing skills and reading skills (3.2, 2.7, and 2.6). Ninety-eight percent of the students thought that
their language skills were better or much better than before (19% said “much better” and 79% said “better”). Sixty-five percent of the students indicated that the online learning environment was more or significantly more helpful than a conventional textbook.

**DISCUSSION**

On average, students’ language proficiency progressed over the course of the year. The c-test scores of the study are higher than the scores achieved in the validation study from the previous year. We found a rather large group of negative gain scores relating to the composition and to the conversational interactions. We found no negative gain scores in the c-test. There are several explanations which may account for these negative scores. If these scores are correct and not an artifact of the grading, it would indicate that as students start to create more with the language as they do in open-ended tasks, they may employ more difficult vocabulary and constructions in the language and may be less successful in using these constructions effectively. We would therefore expect gain scores in open-ended tasks to dip initially, indicating a u-shaped learning curve, which is well documented in the literature on learning (e.g., Karmiloff-Smith, 1992). A u-shaped learning curve has also been documented for learning German as L2 (e.g., Siebert-Ott, 2000). Furthermore, writing and speaking tasks are fundamentally different and have different processing requirements than less open-ended tasks such as a c-test. When writing or speaking, learners need to pay attention to the language on several levels at the same time and may therefore be under processing constraints. The production of accurate meaning to ensure successful communication may dominate syntactic or morphological output.

Assessment using global rating scales across various proficiency levels may differ as well. In similar studies, instructors paid more attention to vocabulary and pronunciation when evaluating a speaker at a lower level. The importance of fluency and grammar became more important in judging as the proficiency level increased (Iwashita, 2010).

The majority of students quit the grammar work only after they had achieved a relatively high score, namely 90% to 100%. This result surprised us as students were told to use the grammar as a resource to learn new vocabulary and practice new grammar points and that we would give them credit for being engaged with the material not for achieving a high score (students could take the exercises as many times as they wanted). This may indicate that receiving feedback on the quality of their work may be beneficial for the length of time a student is engaged as achieving a high score was a priority. Some students used the online grammar primarily during the A.M. period and some primarily during the P.M. period. Students who worked with the grammar primarily during the P.M. hours consistently outperformed the A.M. students. At this point we have no explanation for the result, only that the result may be a consequence of a self-selection.

We did not find any significant difference between usage pattern of the online grammar and the results of the written composition, the oral chat, or the c-test. However, we found a highly significant correlation between usage pattern of the
online grammar and the quiz grades. The better the scores on the grammar exercises, the better the quiz grades.

The online grammar includes a very balanced approach as it affords practice in all four modalities and from controlled to semi-controlled to open ended activities and we were expecting a more uniform increase in proficiency across the three assessment types. Quizzes were always given at the conclusion of each grammar chapter. The quizzes contained controlled, semi-controlled, and open-ended test parts. However, each task was designed to focus on the grammar point(s) in question, even in more open-ended short paragraph writings. Therefore, performance on quizzes may not represent a true mastery of certain linguistic features but may only represent the recall of an item stored in short-term memory. This is also corroborated in the literature where certain grammar items may not be “learnable” through online practice or any extrinsic intervention until an appropriate development stage is reached. These items may simply follow a developmental sequence that cannot be altered through intervention (i.e., Macaro & Masterman, 2006; VanPatten, 2011).

Limitations of the Study

One limitation of the study is the lack of a control group. We had only a certain number of subjects who participated in all of the assessments. Splitting up these students into a technology-based section and a regular-class section would not have given us enough subjects in each group to make statistically meaningful comparisons. We also wanted to maintain experimental conditions as rigorously as possible without compromising the integrity of the educational situation: it would have been difficult to deny the use of technology to some students but not to others.

To assess proficiency of the subjects pre and post, we used a variety of assessment tools. The use of human judged results in grading the compositions and the oral chats is always problematic because of the subjective nature of the evaluation. To minimize the effect, we organized group evaluation practice sessions with the instructors and we performed an inter-rater reliability test. These procedures resulted in a high rate of inter-rater reliability.

CONCLUSION

As indicated in the introduction, the main focus of this study was to explore computer mediated grammar practice and its effect on language learning over time with respect to various tasks. The results of the study suggest that the online grammar practice integrated in a blended learning environment was beneficial to language learning. On average, students’ language proficiency improved over the course of the year as seen in the pre- and post-test results for the c-test, the composition, and the oral chat. However, in the absence of a control group, the impact of the online grammar practice as opposed to other factors could not be independently measured.

Evaluating the use of the online grammar with regard to learning revealed sev-
eral important lessons. Students were engaged users as seventy-four percent said they used the online grammar every day or every other day. They found the interactive exercises very helpful for practicing grammar points and thought that their grammar skills improved. Online tracking showed that students paid attention to how well they did in each practice exercise as they repeated an exercise until they received a fairly high score. The data further showed that prolonged online grammar practice correlated with higher gain scores on chapter quizzes. However, prolonged online grammar practice for many did not have a measurable effect on open-ended tasks such as a composition or an oral chat.

The analysis of the three proficiency tasks revealed some unexpected results. We found a large set of negative gains in the composition, in the oral chat results, but not in the c-test. This may have implications for testing in general. Since the gain in c-test scores appears to be linear (this was seen in both c-test studies), this type of testing may be a good assessment tool for placement purposes or for language assessment in computer-mediated language environments. Eckes and Grotjahn (2006) have previously shown that scores from carefully developed c-tests in German can be very reliable and can be used as an efficient and economical means for testing writing and oral proficiency.

The analysis of student usage and performance data has also given us some ideas on the development of effective multimedia materials. The results of this limited study suggest that it may be beneficial for students to extend their grammar practice to periods outside of class using an online environment. This is a first step. Further exploration in designing a better learning environment (a better interface with more varied feedback, more meaning-focused exercises that allow students to notice and practice target grammatical forms in context, and open-ended production tasks that take advantage of the newest technology such as face-time interactions) may show that prolonged online practice can promote the use of correct grammar in quiz-type assessments as well as in more open-ended spontaneous tasks.

NOTES

1 This project was part of the “Cost-effective Uses of Technology in Teaching” initiative funded by the Andrew W. Mellon Foundation and the Hewlett Grant for Teaching Excellence at Northwestern University. I would like to thank the staff at the Searle Center for Teaching Excellence, the Multimedia Learning Center in the Weinberg College of Arts and Sciences, and the instructors in the Department of German for supporting this research.

2 We had worked with Libra extensively creating various lessons for an online video-based project also part of the material used in the intermediate German program (Lys, 1997; Wittig, 1999).

3 Arras, Eckes, & Grotjahn (2002) for example have shown a high correlation between the c-test and oral proficiency measured by a Simulated Oral Proficiency Interview (SOPI).
4 This validates Klein-Braley (1985, p. 84) who states that “C-Test scores will become successively higher as the subjects become more proficient in the language” (cited in Mahmood, 2008).

5 Previous oral proficiency testing we had done with our second-year students showed that 95% of the students were in the intermediate mid range and we doubted that this scale would be refined enough to show gain over a six to eight months period.

6 It is not clear whether this is an artifact of the grading (the human-judged scores) or whether it shows us the fundamental difference between the tasks. The validity-based explanation of negative gain scores was tested by replacing negative gain scores with 0 gain; positive scores remained the same. It was demonstrated that the inference about underlying ability change would be the same. This points us into the direction of the type of task used as a catalyst for the observed differences in gain scores.

REFERENCES


APPENDIX A

SAMPLE UNIT IN INTERMATIK

Cultural Section (Kultur)

Part 1: About the Topic (Zum Thema)
A set of personal questions to engage students in the cultural topic.

Part 2: Researching Facts about the Cultural Topic (Die Kultur erforschen)
A set of research questions pertaining to cultural information. Students are directed to various Internet sites to find the answers. These tasks are designed to improve cultural knowledge.

Part 3: Things to Discuss (Zum Diskutieren)
A broader (sometimes controversial) discussion question based on the cultural topic.

Part 4: Extra (Fakultativ)
A small set of additional cultural tasks for small group work or class discussion.

Vocabulary Section (Vokabular)

Part 1: New Vocabulary (Neue Vokabeln)
A list of vocabulary items relevant to the cultural topic.

Part 2: Vocabulary Practice (Vokabelübungen)
A set of exercises to practice the new words and expressions. The first is usually a drag-and-drop exercise to help students recognize the English and German equivalents. The remaining exercises are designed to practice and solidify the new expressions in context.

Grammar Section (Grammatik)

Part 1 through Part 6: Grammar Presentations and Practice (Grammatik lernen)
These parts contain presentations of special grammar points (such as, for example, reflexive verbs, time expressions with the dative case, etc.). The explanations are in English with German examples including sound files for pronunciation. Each grammar point is accompanied with a series of exercises to practice the grammar points in context. Exercise formats are multiple-choice questions, check list questions, pull-down menus, open-ended responses, and various drag-and-drop options.

Review (Wiederholung)

Part 1 … Part 4: How much did I learn (Was kann ich jetzt)
These four parts test the new grammar points through a variety of exercises and tasks. The exercises follow a similar format as the exercises in the grammar part.

Part 5: Listening comprehension (Zum Zuhören)
The exercise in this part is a global listening comprehension task.

**Part 6: Free writing section (Zum Schreiben)**

The exercise in this part contains a video clip to elicit open-ended responses for a more creative use of the language integrating all of the grammar points practiced.

**Resource Section (Ressourcen)**

The resource section is includes active vocabulary lists, irregular verbs, grammar tables, as well as site explanations and technical help.

---

**APPENDIX B**

*Example of C-Test used in Assessment of General Language Proficiency and Grammar*

Bin ich etwa ein Baby?
Auch Teenager in Deutschland haben Ärger mit den Eltern. Teenager haben manc_______das Gef_______, dass ih_______ Eltern s_______ noch w_______ ein Kleinki_______ behandeln, obw_______ sie sc_______ fast erw_______ sind. Die mei_______ Eltern sa_______, w_______ der Jugen_______ tun so_______, und las_______ ihn (oder sie) ni_______ selber entsc_______: Da hi_______ nur eins: man mu_______ mit d_______ Eltern re_______, u_______ sie zu überz_______, dass sie ei_______ vertrauen kön_______: Dann haben auch die Eltern eine Chance zu erklären, was ihnen nicht passt.
APPENDIX C

Sample Discussion Topics for Oral Assessment

The Drinking Age in the USA
Young people in Europe can drink alcohol at the age of sixteen. In most states in the USA, one has to be twenty-one years old in order to drink legally. Do you think that the drinking age of twenty-one is a good idea for young people in the USA or should it be lower? Please explain. Describe how Northwestern students feel about alcohol. Do many students drink? Why and when? Please describe any relevant experiences that you have had or that you have witnessed.

Do Americans have a healthy lifestyle?
Many people in Europe believe that Americans do not lead healthy lives. Do you agree or disagree? How would you describe the average lifestyle of Americans? Do you think student life at universities encourages an unhealthy lifestyle? What are some things that you personally do or have done to stay healthy and fit?

Living Situations and Housing at American Universities
In Germany, students entering the university are not required to live on campus. Most often they rent their own apartment. How is the situation here in the USA? Are students required to live on campus for part of their university experience? Why would a student rent an apartment in Evanston? Describe your own living situation at Northwestern. What are the advantages and the disadvantages? Please describe any relevant experiences that you have had.

APPENDIX D

Sample Topic for Written Assessment

Instead of working in the summer, you have decided to take off for Europe and study German. The problem is, however, that you have not yet told your parents of your decision... En route to Frankfurt via Lufthansa, you take a piece of airline stationery and pen a letter to them: first, to let them know of your decision, secondly, to tell them how this decision came about, and thirdly, to let them know what you’re doing or planning to do, so that they don’t worry. (Naturally, you write in German in order to really impress your folks...) You may want to tell them where you’re going; how you plan to travel (will you hitchhike?); what you’re taking along; how long you are going to stay; in what cities (universities) you plan to study; and when you plan to return.