Abstract
This study investigates the effectiveness of computer-mediated instruction in a Greek State Lykeio comparing CALL learning outcomes with the outcomes of traditional teaching.

Most of the studies on CALL that have been carried out so far focus on CALL’s holistic nature and they are based on analyses of outcomes in on-line teaching courses or motivation studies. On the other hand, comparatively little research has been conducted on CALL’s effectiveness in teaching isolated linguistic input. This study compares the learning outcomes of teaching the same isolated form (inversion) to two groups of randomly selected students in a Greek State Lykeio. Additionally, the study looks onto potential differences in CALL effectiveness between males and females.

The study looks into the relevant literature and presents the work that has been carried out so far with reference to similar studies that compare learning outcomes between CALL and traditional classroom environments.

A pretest-posttest randomized experimental design is implemented in order to study and compare the learning outcomes of the two groups as well as the findings concerning males and females separately.

The study found statistically significant differences favouring the traditional teaching methods, whereas no statistically significant differences were noted in the performance between males and females.

Even tough the differences could be attributed to the mentality of the students in a Greek State Lykeio, in order to come to more definitive conclusions, a study comparing CALL’s learning outcomes when teaching various forms together with an investigation of the retention factor is required.
Περίληψη

Η παρούσα μελέτη διερευνά την αποτελεσματικότητα της διδασκαλίας της αγγλικής γλώσσας μέσω της πληροφορικής σε ένα Ελληνικό Δημόσιο Λύκειο συγκρίνοντας τα μαθησιακά αποτελέσματα της με αυτά της παραδοσιακής διδακτικής μεθόδου.

Οι περισσότερες μελέτες πάνω στην διδακτική μέσω της πληροφορικής εστιάζονται στην χρήση των υπολογιστών ως κυρίως διδακτικό μέσο. Βασίζοντας δε, σε αναλύσεις αποτελεσμάτων μετά από παρακολούθηση μαθημάτων μέσω του διαδικτύου ή πάνω σε μελέτες που αφορούν διαφοροποιήσεις στα κίνητρα των μαθητών όταν διδάσκονται μέσω της πληροφορικής.

Αντιθέτως, λίγη σχετικά έρευνα έχει διεξαχθεί η οποία να αναφέρεται στην αποτελεσματικότητα της πληροφορικής ως διδακτικής μεθόδου για μεμονωμένα συντακτικά ή γραμματικά φαινόμενα. Η παρούσα μελέτη συγκρίνει τα μαθησιακά αποτελέσματα της διδασκαλίας ενός μεμονωμένου φαινομένου (αντιστροφή) σε δύο ομάδες τυχαία επιλεγμένων μαθητών σε ένα Ελληνικό Δημόσιο Λύκειο. Επιπροσθέτως, διερευνώνται οι πιθανές διαφορές στην αποτελεσματικότητα της πληροφορικής ως διδακτικού μέσου μεταξύ αρρένων και θηλέων.

Στην μελέτη αυτή παρουσιάζεται η σχετική με το αντικείμενο βιβλιογραφία καθώς και αντίστοιχες έρευνες οι οποίες συγκρίνουν τα μαθησιακά αποτελέσματα της πληροφορικής ως διδακτικού μέσου με τα αποτελέσματα της παραδοσιακής διδακτικής.

Για την εκπόνηση της μελέτης και την σύγκριση των αποτελεσμάτων μεταξύ των δύο διδακτικών μεθόδων καθώς και την σύγκριση των αποτελεσμάτων μεταξύ των δύο φύλων, επελέγη ο πιεραματικός σχεδιασμός τυχαίου δείγματος με μετρήσεις πριν και μετά το πείραμα.

Η έρευνα κατέδειξε στατιστικά σημαντικές διαφορές υπέρ της παραδοσιακής διδακτικής, ενώ δεν ευρέθησαν στατιστικά σημαντικές διαφορές στην αποτελεσματικότητα των δύο μεθόδων μεταξύ των δύο φύλων.

Αν και οι διαφορές που ευρέθησαν θα μπορούσαν να αποδοθούν στην νοοτροπία των μαθητών του Ελληνικού Δημοσίου Λυκείου, για να καταδειχθούν οριστικά συμπεράσματα, απαιτείται μία μελέτη η οποία να διερευνά τα μαθησιακά αποτελέσματα της πληροφορικής στην διδασκαλία της αγγλικής γλώσσας σε διάφορους γραμματικούς και συντακτικούς τόπους, λαμβάνοντας ταυτοχρόνως υπ’ όψιν τον παράγοντα της χρονικής διάρκειας συντηρήσεως των φαινομένων στην μνήμη των μαθητών.
Acknowledgements

The author wishes to acknowledge his appreciation to the professionals and family who supported the creation of this study.

Specifically, the author would like to express gratitude to his advisor, Dr. A. Taglidis, for providing mentoring, inspiration, and constant, enthusiastic support.

In addition, the author would like to acknowledge the support and encouragement of his wife, Dora Simpsa-Hartzoulaki.
List of Tables

TABLE 1: GENERAL DESCRIPTIVE STATISTICS ................................................................. - 23 -
TABLE 2: CLASS DESCRIPTIVE STATISTICS ................................................................. - 24 -
TABLE 3: FINAL SHORTLIST DESCRIPTIVE STATISTICS .............................................. - 25 -
TABLE 4: PRETEST DESCRIPTIVE STATISTICS – OVERALL .......................................... - 26 -
TABLE 5: PRETEST DESCRIPTIVE STATISTICS – FINAL LIST ....................................... - 27 -
TABLE 6: PRETEST, POSTTEST DESCRIPTIVE STATISTICS FOR ALL STUDENTS ....... - 28 -
TABLE 7: MEAN, MEDIAN AND MODE FOR ALL STUDENTS ......................................... - 29 -
TABLE 8: PAIRED SAMPLES T-TEST ............................................................................ - 30 -
TABLE 9: POSTTEST DESCRIPTIVE STATISTICS ......................................................... - 31 -
TABLE 10: GAIN DESCRIPTIVE STATISTICS ............................................................... - 32 -
TABLE 11: ANCOVA FINDINGS FOR THE TEACHING METHODS ............................... - 33 -
TABLE 12: DESCRIPTIVE STATISTICS FOR THE GAIN MEANS BETWEEN MALES AND FEMALES - 34 -
TABLE 13: ANOVA FOR ‘SEX’ ................................................................................... - 35 -
TABLE 14: ANOVA FOR THE TEACHING METHOD FOR MALES AND FEMALES .......... - 36 -

List of Figures

FIGURE 1: PRETEST, POSTTEST MEANS OF THE GROUPS ......................................... - 37 -
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Περίληψη</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>iv</td>
</tr>
<tr>
<td>Chapter I – Introduction</td>
<td>- 3 -</td>
</tr>
<tr>
<td>Historical Foundations of CALL</td>
<td>- 4 -</td>
</tr>
<tr>
<td>Research Question</td>
<td>- 6 -</td>
</tr>
<tr>
<td>Subsidiary Question</td>
<td>- 6 -</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>- 6 -</td>
</tr>
<tr>
<td>Need for the Study</td>
<td>- 7 -</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>- 8 -</td>
</tr>
<tr>
<td>Assumptions</td>
<td>- 8 -</td>
</tr>
<tr>
<td>Limitations</td>
<td>- 8 -</td>
</tr>
<tr>
<td>Organizational Format Followed</td>
<td>- 8 -</td>
</tr>
<tr>
<td>Chapter II – Review of Literature</td>
<td>- 10 -</td>
</tr>
<tr>
<td>Educational Technology Background</td>
<td>- 11 -</td>
</tr>
<tr>
<td>Evaluating CALL</td>
<td>- 13 -</td>
</tr>
<tr>
<td>Recent Comparative Studies</td>
<td>- 14 -</td>
</tr>
<tr>
<td>Focus on Form</td>
<td>- 15 -</td>
</tr>
<tr>
<td>Developing CALL Tasks</td>
<td>- 17 -</td>
</tr>
<tr>
<td>Animation</td>
<td>- 19 -</td>
</tr>
<tr>
<td>Research Methods</td>
<td>- 19 -</td>
</tr>
<tr>
<td>Conclusion</td>
<td>- 20 -</td>
</tr>
<tr>
<td>Chapter III - Methodology</td>
<td>- 21 -</td>
</tr>
<tr>
<td>Experiment Design</td>
<td>- 22 -</td>
</tr>
<tr>
<td>Population</td>
<td>- 22 -</td>
</tr>
<tr>
<td>Selecting a Homogeneous Group</td>
<td>- 23 -</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>- 25 -</td>
</tr>
<tr>
<td>Piloting and Pretesting</td>
<td>- 26 -</td>
</tr>
<tr>
<td>Random Selection</td>
<td>- 26 -</td>
</tr>
<tr>
<td>Random Assignment</td>
<td>- 27 -</td>
</tr>
<tr>
<td>Software Design</td>
<td>- 28 -</td>
</tr>
</tbody>
</table>
Chapter I – Introduction
Historical Foundations of CALL

In recent years, there has been widespread interest in using computers to enhance learning processes. This is due to the rapid growth of computer technology, which led to more sophisticated, more user-friendly, more manageable in terms of size and weight, and much cheaper machines.

Computers have become so widely spread, that what originally was highly sophisticated, extremely expensive equipment meant for military purposes, has now become a household name. Communication between computers has also become faster, more reliable and more easily accessible in terms of hardware and software. Today, thirty-three years after the first network was set up—according to Cerf (2002) it was Labor Day in 1969—and twelve years after the World Wide Web (WWW) was introduced (Connolly, 2000), communication through computers—in other words, the Internet—is a fact of life.

This development has led foreign language teachers to try to find ways of incorporating computers and information technology in general into the educational processes, especially in cases when the language taught is not the students’ mother tongue (Virvou, Maras & Tsiriga, 2000). Yet, simply introducing computers into the language classroom for the sake of taking advantage of technology is not enough. Meloni (2002), for example, suggests that technology is to be used only when it offers something that conventional classrooms cannot. Consequently, one might say that researchers and educators must investigate those features of Computer Assisted Language Learning (CALL) tasks that do promote learning. This question attracted researchers’ attention, who since the 1960s, have been trying to investigate the best ways of using computer technology for instruction across subject areas and to design valid ways of evaluating its effectiveness (Chapelle, 2001a:28). As a result, quite a large number of studies that compare CALL with traditional classroom learning have been carried out. Relative literature research reveals that the vast majority of those comparative studies investigate the holistic nature of CALL particular to higher education (Chapelle, 2001a:29), e.g. on-line teaching courses or motivation studies, and as Chapelle puts it, “surprisingly little research” has been carried out on the effectiveness of CALL when teaching specific aspects of the language (2001a:69).

The integration of CALL into educational processes in Greek Secondary Schools has been a goal for the Greek Ministry of Education and an expectation for the educational community for quite some time. In the last two years, there has been a major development in the quality and quantity of technological equipment in the Greek State Secondary Education. The majority of schools (1350) have been equipped with computer labs with access to the Internet (Edunet, 2002),
opening new possibilities for teachers and learners alike. Therefore, teachers in Greek State Secondary Education schools are faced with the challenge of devising ways to integrate CALL into their own syllabus. This is a difficult task, since CALL is certainly an innovation for Greek Lykeia, let alone the fact that not all teachers are trained to use computers.

The very detailed syllabus that the Hellenic Ministry of Education (Hellenic Government Gazette issue 2, 1999) has issued encourages the use of computers in the teaching of English, however without specifying ways of doing so. What is specified is a detailed list of forms, functions, and structures that are to be taught throughout the three year span of the Hellenic Secondary Education. The teachers are, in that way, guided towards covering the material set -the specified forms and functions- in the best way they see fit for their students.

The term *form* has been given a number of definitions. Doughty and Williams (1998:4) clarified the issue by differentiating the term *form* from the term *formS*. They adopted Long’s (1988) suggestions who defined isolated linguistic features (e.g. grammar) as *formS* and the overall engagement in meaning as *form*. Doughty and Williams (1998) stressed that the two terms are not polar opposites; instead, they are interwoven in the sense that form entails isolation or extraction of linguistic features from context, whereas *formS* is limited to isolated features. This approach has been adopted by the author throughout this study.

A question that is raised here is whether these isolated features of the language (*formS*) could be taught effectively within the framework of CALL. Motteram (2000a) and Kenning & Kenning (1990) see CALL as an add-on, “making a greater contribution to marginal rather than central elements” (Kenning & Kenning, 1990:90) of a language curriculum. Whenever computer-based activities were implemented in classroom teaching, they were seen as supportive of the curriculum and integration of technology into the curriculum was the driving force leading to learning enhancement (Motteram, 2000a). Motteram quoting the electronic journal *From Now On* says that technology in foreign language learning is particularly suited to support scientific enquiries and engagement in collaborative activities with the teacher acting as orchestrator (2000a:17). Evidently, what Motteram and Kenning & Kenning imply is that technology is suited to support teaching *form*. This, leads to the aforementioned question of whether CALL could be equally effective when used to teach *formS*. 
Research Question
This study is structured on a framework encompassing the various parameters that hold a part -direct or indirect- in answering the following question:

• Is computer-mediated instruction more, less than or equally effective as traditional teaching in a Greek State Lykeio?

The study compares the learning outcomes achieved through computer-assisted instruction with those obtained in a traditional classroom when teaching one of the forms -in this case, inversion- to two groups of randomly selected students in the second and third grade of Lykeio.

Subsidiary Question
Additionally, the following supplementary question is explored:

• Is computer-mediated instruction equally effective for males and females?

Purpose of the Study
This study explores the effectiveness of CALL when teaching specific grammatical items, in comparison with employing conventional methods of teaching, in a Greek State Lykeio. By conventional methods of teaching, the author is referring to the traditional class where the teacher uses the blackboard as the visual stimulus and face-to-face interaction for the development of the lesson, whereas by computer-mediated instruction, the author is referring to a fully equipped multimedia computer lab where students work on their own.

This study is based upon a pretest-posttest randomized experimental design (Trochim, 2002) with the objective of comparing CALL’s effectiveness in teaching forms with the effectiveness of traditional teaching, by analyzing the outcomes of a computer-based instruction with those of a traditional class. The form that was selected was ‘inversion’, as it is one of the points required by the syllabus for Lykeio. For the needs of the experiment, a website was designed and uploaded on the Internet. The treatment group would discover the rules for the selected form by means of visiting the website and navigating through its web pages that presented the various cases of inversion, whereas the control group would be taught the same form in a traditional classroom environment without employing CALL whatsoever.
Need for the Study

The early days of CALL (from the beginning to the mid 1980s) were marked by an aura of ‘conscious learning’. CALL activities were designed aiming at learning the language through drills and exercises focusing on the mechanics of the language rather than on its communicative value. However, this period saw the height of Krashen’s popularity, who presented subconscious ‘acquisition’ as a different process from conscious ‘learning’ and argued that the former is the most effective (Chapelle, 2001a). This resulted in a trend to invent CALL that would promote ‘acquisition’ rather than ‘conscious learning’. There were different views as to the ability of computers to promote either acquisition or conscious learning. Higgins & John (1984) and Underwood (1984) denounced the link between CALL and explicit teaching, whereas Abraham (1985) found that analytic ESL learners performed better on posttests when they had used a rule presentation approach, and holistic learners performed better after using software presenting examples of the structure. Newer studies merge the two theories and agree that learners need to notice and attend to linguistic form in order to acquire the language (Schmidt, 1990; Robinson, 1995). The introduction of graphics, sound, animation, and video –all in one machine- and the possibility of exchanging these pieces of information through the Internet led to the design of CALL activities of a more communicative value as situation specific environments of authentic nature were made possible (Bachman & Palmer, 1996; Motteram, 2000b). Therefore, whenever these days one speaks of various applications of computers in language learning, one usually refers to the various ways of communicating via the Internet and searching for materials in it (sound, video, texts) in order to compile them into a full project (Murray, 2000; Motteram, 2000b). Still, the “traditional” CALL function of teaching form\$ with the aid of computers “...seems to be in the periphery of research” (Tarnopolsky, 2001).

So, how effective might computer-mediated explicit instruction be, compared to explicit instruction in a traditional classroom environment? This question can be viewed from different perspectives depending on the teaching situation and teaching objectives. Russell (2002) set up a website where he has gathered a large number of comparative studies between CALL and traditional classrooms, some of which end up favouring CALL, others favouring traditional classrooms, but the largest number finds ‘no significant difference’ between the two. The last 54 studies (from 1998 to 2002) address universities and online versus distance learning, but there are no studies on the effectiveness of CALL compared to traditional classrooms when teaching form\$. For the aforementioned reasons, the author found it necessary to carry out an experiment investigating this fine point.
Definition of Terms

- Lykeio: Greek Secondary Education Institution addressing adolescents from 15 to 17 years of age.
- Traditional Class: Conventional term for the class where no computers are used in the process of the lesson, the visual stimulus is the teacher and the blackboard, and the interaction patterns are between the teacher and the students, and between the students themselves.
- Form: Overall engagement in meaning, encompassing elements of language studied in context.
- FormS: Isolated linguistic features, formal elements of language extracted from context.

Assumptions

- The teacher carrying out the lesson is not a novice or inexperienced. He does not favour any of the teaching methods and the results of the study are not affected by his way of teaching.
- The students do not show a negative attitude towards technology and cannot be considered to be technophobic.

Limitations

The number of the students who were sampled for the experiment that is presented in the study was quite small in proportion to the population the author wanted to generalize to. Furthermore, the experiment was conducted within the framework of the regular teaching hours in the author’s school; therefore, the sample is also limited in terms of the geographical area covered.

Organizational Format Followed

The first chapter constitutes an overall review of the context around which the study revolves, sets the study objectives, and presents background information relative to the study. The second chapter presents an extended review of the literature pertaining to CALL and its various uses in the foreign language classroom, with reference to similar studies that compare -at various levels- learning outcomes between CALL and traditional classroom environments. The third chapter describes the methodology, the procedures and practices that were followed while carrying out the experiment that sets the foundation for this study. The following chapter discusses the study’s findings and results. It presents the descriptive and inferential statistics for the experiment and analyses the extent to which the findings could be generalized to the
representative population. In the fifth chapter, the author discusses the results that were presented in the previous chapter, focuses on their implications, and makes some recommendations that might be considered by language teaching professionals when integrating CALL in the foreign language learning processes.
Chapter II – Review of Literature
**Educational Technology Background**

The issue of devising ways of making the most out of computer technology to the benefit of the foreign language learner has been the concern of researchers since the 1960s when a number of projects were undertaken –mostly by individuals who had access to computer labs in universities- to explore how the computer could be used for foreign language teaching in higher education. Since then, numerous journals devoted to the topic of educational technology have appeared, contributing to the educational technology literature and turning CALL into a distinct professional community (Chapelle, 2001a).

The early days of CALL were influenced by Skinner’s ‘programmed learning’, which embraces specific practices consistent with the ‘systems approach’ to instructional design (Chapelle, 2001a:28). Skinner’s ‘programmed learning’ breaks learning up into its components, which are described as observable behaviours defined in cognitive terms. Wilkins (1976) termed this approach ‘synthetic’ and described the synthetic syllabi as those in which “…parts of the language are taught separately and step by step so that acquisition is a process of gradual accumulation of parts until the whole structure of language has been built up…At any one time the learner is being exposed to a deliberately limited sample of language” (Wilkins, 1976:2).

When Long & Robinson (1998) review approaches to language teaching down the years, they say that the approach of analyzing the target language into words, grammar rules, structures and patterns, has been a long-standing popular position and they add that synthetic syllabi are still used in the vast majority of classrooms all over the world (1998:15).

This approach to instruction is also followed by Jones & Fortescue (1987) who give a picture of what they call “traditional picture of computer-assisted language learning” saying: “It is still common to find CALL described as a means of ‘presenting, reinforcing and testing’ particular language items” (Jones & Fortescue, 1987:5). However, later they reject this picture of CALL and characterize it ‘unfortunate and misleading’ because it implies isolation of the learner and the machine from the communicative classroom context, and because it leads to a development of the lesson per se, without any consideration of the teacher’s methodological processes.

Instead, they present computers as a flexible classroom aid linked with ordinary classroom work. Yet they do not stray away from the trends of the day, as one of the roles they see for computers is “the computer as knower-of-the-right-answer” where computers are used to present particular language features e.g. grammar, vocabulary, reading comprehension and listening comprehension (1987:5). The novelty comes when they introduce the communicative nature of computers in the role of “stimulus” and they stress that one of the uses of computers is “...to
provide learners with something to talk about and to write about” (Jones & Fortescue, 1987:6). They also find a role for the teacher, who is responsible for selecting the area of language s/he will place the most emphasis on at any one time. They say that the teacher should tailor the content of CALL activities to the learner’s needs “be they grammatical, lexical or text-based” and computer work should not be an end in itself. Instead, the more CALL is integrated with normal classroom work, the more relevant it will be (1987:101).

This early 80s shift in CALL methods was also voiced in the works of Higgins & John who rejected the notion that computers serve only the conscious process of learning and do nothing to facilitate acquisition (1984:17). Following the same concept, Underwood said that computers are not limited to promoting learning from drills and tutorials, but the only limitation in CALL is the way computers are used (1984:50). This initiated a swinging towards acquisition rather than structured learning and set the foundations of ‘communicative CALL’ which was consistent with Krashen’s perceptions for subconscious ‘acquisition’ as opposed to the less effective –according to him- conscious ‘learning’ (Krashen, 1982). Warschauer & Healey (1998) summarize this period’s dominant philosophy listing four key directions towards which CALL activities should be oriented; they said that CALL activities should

- Focus more on using forms than on the forms themselves
- Teach grammar implicitly rather than explicitly
- Encourage students to generate original utterances rather than use prefabricated language
- Encourage students to use the target language predominantly or even exclusively.

Following the development of CALL theories down the years, Warschauer & Healey (1998) saw a parallel evolution between computer technology and CALL approaches. They associate the earliest ‘behaviouristic CALL’ with the mainframe technology, the ‘communicative CALL’ with the PC technology and the latest ‘integrative CALL’ with the multimedia networked computer technology. This latest approach incorporates the task-based, project-based and content-based approaches and seeks to integrate language learning and language use in authentic environments. Warschauer & Healey (1998) see that “many of the changes in CALL paradigms flow from economic and social changes” and they add that the wealth of information available to students through networked computers makes the need for memorization less important than the need for developing effective search strategies in order to locate and retrieve the information one wants. Their ‘integrative CALL’ approach sees language learning as using the variety of technological
tools available in an on-going process of learning rather than visiting the computer lab for isolated exercises. This approach simulates real life as “...communicating via computer has become an essential feature of modern life in the developed world” (Warschauer & Healey, 1998). Nevertheless, even in the ‘integrative CALL’ approach, Warschauer & Healey find a place for drilling and practicing forms albeit as a small part of language learning.

Evaluating CALL
Effective evaluation of CALL outcomes has presented difficulties, and opened new grounds for discussion among CALL analysts. The first articles on CALL evaluation appeared at the beginning of the 1980s, when CALL started becoming widely known and drew researchers’ attention to itself. Summarizing the conclusions drawn from the PLATO project (Programmed Logic for Automatic Teaching Operations) in the 1970s, Hart says that developers of CALL materials “have given far too little attention to evaluation” (1981:16). He also acknowledges the difficulties in devising sound methods of evaluating CALL and he suggests researchers should present clear arguments why that is so and provide alternative analyses. An analysis of the difficulties inherent in evaluating CALL has been given by researchers (e.g. Chapelle, 2001a; Pederson, 1987), who focus on the incompatibility of evaluation methods between CALL and traditional classroom environments. Chapelle (2001b) says that knowledge about classroom teaching probably does not apply directly to CALL because of the different configuration of variables such as learners’ and teachers’ roles, location, and interactivity. Researchers have developed diverse theories on how CALL should be evaluated, ranging from views that research should adopt an experimental or quasi-experimental design (Alderman, 1978; Kulik, Kulik & Schwalb, 1986) to the perspective that such research would be too narrowly focused and product-oriented (Clark, 1985; Papert, 1987). Pederson (1987) goes even further by saying that comparisons of CALL versus classroom learning outcomes create an irony, because a sophisticated modern educational tool is investigated through crude, outdated educational research methods.

The need to evaluate CALL emerged from the need to inform major funding institutions about the effectiveness of the computer as a tool for teaching academic material. This need dates back to the 1960s when several universities initiated large-scale projects to investigate CALL’s effectiveness. Despite the criticism on the evaluation methods (e.g. Pederson, 1987; Clark, 1985; Papert, 1987), the reference point throughout most of these projects was learning outcomes in a traditional classroom; therefore, “the objective was to compare learning outcomes achieved through computer-assisted instruction with those obtained in traditional classrooms” (Chapelle,
Chapelle summarizes the findings from these early projects saying that there are studies “...in which CALL users performed better than learners who did not use CALL, others in which no differences were found, and even a few in which the control group performed better” (Chapelle, 2001a:30). Based on this, she comments that no conclusions can be drawn about the computer as a teaching method. However, the findings from these projects urged researchers to set up further investigations comparing CALL with traditional class outcomes.

**Recent Comparative Studies**

Since the 1960s, quite a large number of studies investigating CALL’s effectiveness have been conducted, mostly based on the same concept – comparing CALL outcomes with traditional classroom outcomes. Much of the literature related to such studies is collected on Russell's (2002) “No Significant Difference” website. The introductory page says that the site comprises entries from 355 comparative studies, which are published in the form of research reports, summaries and papers. Seventy-three of the 78 studies conducted in the last seven years focus on the effectiveness of distance education projects run by higher education institutions as opposed to traditional classroom teaching in those same institutions. The majority of those studies were conducted by American universities that offered their distance education courses in an online format, requiring their students to take advantage of the networked facilities of the institution. Thus, by 'distance education projects', online teaching is implied in which the students access the required material through the institution's network. The remaining five studies focus on the effectiveness of specific courseware, developed by the institutions that carried out the corresponding study. They investigate the effectiveness of the courseware in question, in comparison with traditional classroom teaching. The studies have typically been used by distance educators to defend the quality of their courses and programs against the view that learning takes place only in a physical classroom (Twiggs, 2002). Therefore, they measure students’ performance through posttest scores and compare the results with the performance of students who completed the equivalent on-campus course (e.g. Johnson, 2002; Colorito, 2001; Nesler & Hanner, 2001; Carey, 2001, Carnevale, 2001). Although these studies investigate CALL effectiveness, they differ from the one the author is conducting, as they investigate outcomes from whole courses (a type of evaluation that can be called macro-evaluation), whereas the author's study looks into the effectiveness of technology when teaching a discreet point.

A review of the entries published in the Russell website for the last seven years, shows that of the 78 published reports, 21 favour technology-based learning, 4 favour traditional classroom teaching, and 53 show no significant difference between the two. The findings are consistent
with Chapelle’s (2001a) findings whose research -based on the works of Chapelle & Jamieson (1989; 1991), Chapelle, Jamieson & Park (1996) and Dunkel (1991), has found that the results of similar comparative studies are almost equally distributed amongst favouring technology, finding no significant differences and favouring traditional teaching. Thus, the results “...do not warrant conclusions to be drawn about ‘the computer’ as a teaching method” (Chapelle, 2001a: 30).

All in all, Tomas Russell's (2002) collection of reviews appears to support the argument that when distance learning and traditional classroom environments are compared, no significant difference in student performance is noted. However, a number of researchers (e.g. Long & Javidi, 2001; Jonassen, 1985) attribute these findings to differences in the independent samples' characteristics. Threlkeld & Brzoska (1994) have found that distance learning subject samples-in most of the studies- were drawn from individuals who bear very little resemblance to full-time student population. They describe the distance learning subject samples as individuals who were part-time students, preferred not to commute, were self-directed, older than most full-time students, frequently employed full-time, and often married. Hence, the generalization that 'no significant difference exists in learning outcomes between traditional classroom settings and CALL environments', is questionable as it is not based on measures taken from subjects with similar characteristics. It may be true though, when one compares a typical distance/online learner with a typical full-time student in a traditional classroom. Learning outcomes within student groups with the characteristics of full-time students have yet to be investigated.

**Focus on Form**

When researchers investigate the efficiency of CALL versus traditional classroom teaching, a parameter that is to be taken into account is the nature of the learning outcomes that are compared. Most commonly, researchers employ quasi-experimental designs and compare posttest scores measuring the effectiveness of the whole course be it distance learning or face to face. This is a holistic evaluation of learning outcomes, which can give some broad evidence of CALL’s effectiveness but it does not evaluate specific, isolated output. This kind of micro-evaluation has been given little attention, although many researchers stress the importance of focusing on form for acquisition (Schmidt, 1990; Robinson, 1995). Chapelle (2001a) presents the findings of one study (Doughty, 1991) which yielded results favouring a CALL activity that focuses learners’ attention on a particular linguistic form. Oxford, Rivera-Castillo, Feyten, and Nutta (2002) carried out a study addressing adolescents and adults and investigated CALL’s effectiveness from various aspects. They found that technology can focus on grammar and
vocabulary, involve strictly minimal communication restricted to particular forms, and require the learner to contribute only letters or single words. The word ‘forms’ here refers to the isolated linguistic features of a language which contribute to meaning (e.g. grammar, vocabulary etc.). They have been referred to as ‘forms’ (Long, 1988) and have been discriminated from ‘form’ which refers mostly to meaning (Doughty & Williams, 1998), although Doughty & Williams have stressed the fact that “form” and “forms” are not two polar opposites in the way that form and meaning have often been considered to be. They are interwoven features, mutually inclusive, both essential for the extraction of meaning in language. Doughty & Williams (1998) have described a problem that sometimes arises when one needs to discriminate teaching of linguistic forms in isolation from teaching that integrates attention to forms, meaning, and use. So far, “focus on form” has been used for both cases. In order to facilitate the differentiation between form and forms they opted “...to use the contrasting (but not opposite) terms forms-focused instruction and FonF... [an acronym for focus-on-form, which has come into colloquial use]... instruction” (Doughty & Williams, 1998:4). Bruton (2000) offers three alternative realisations when it comes to forms-focused instruction: (1) pre-planned itemized focus through input and/or output (with/without feedback), which corresponds to language-focus tasks, (2) itemized language focus for input and/or output (with/without feedback), which corresponds to skill-focus tasks, and (3) no conscious language focus, which corresponds to open-ended communication tasks. In this study, the term ‘form’ will be referring to isolated features of language as described by Long (1988) and formS-focused tasks will be considered equivalent to Bruton’s (2000) language-focus tasks.

A look into Russell’s (2002) collection of reviewed research studies shows that in the last seven years there is no entry for a comparative study investigating CALL’s effectiveness in teaching forms. The studies –involving CALL focusing on form- which have been carried out so far, either compare one type of CALL materials with another (Doughty, 1991) or they analyze CALL effectiveness in itself (Duquette, Renié & Laurier, 1998; Virvou, Maras & Tsiriga, 2000). Doughty (1991) for example, compares the effects of two different types of CALL activities focusing on relative clauses; Duquette, Renié & Laurier (1998) investigated conditions that promote learning of vocabulary in a multimedia environment and Virvou, Maras & Tsiriga (2000) investigate the effectiveness of an intelligent multimedia tutoring system for the teaching of passive voice of the English grammar for Greek students. In none of the aforementioned studies was there any involvement of traditional teaching. More specifically, in Doughty’s (1991) study, the outcome (learning of relative clauses) was evaluated through a grammatical posttest given to the two experimental groups, and the comparison was between the outcome of
a CALL activity which highlighted relative clauses in a passage and one which did not. Duquette, Renié & Laurier (1998) employed a quasi-experimental design and -using video- identified some of the critical characteristics of input that may help to focus learners’ attention on unknown vocabulary and to remember it. Virvou, Maras & Tsiriga (2000) employed qualitative techniques (questionnaires) in order to evaluate their Passive Voice Tutor and compared its error analysis features with that of 10 human tutors who were asked individually to give their comments on the sample. They considered it successful as its judgment on passive voice errors coincided with that of the human experts, making the ‘tutor’ a potential teacher simulator.

Warschauer & Healey (1998) discriminate four areas of research on CALL. Specifically, they recognize research on:

- the amount and types of interaction between the learners and the computer,
- effects of skill-building software,
- responses to multimedia and
- attitudes towards computers and CALL.

When they look into research on specific skills, they see that it “has tended to focus on student attitudes toward drills and effectiveness of drills” (Warschauer & Healey, 1998) and they add that the results have been mixed due to the types of software, the teacher roles and the diversity of the tasks.

The studies investigating CALL effectiveness when teaching discreet points are comparatively few. They mostly focus on types of CALL materials or skills rather than on grammatical forms. The ones that have been reviewed so far do not lead to definite conclusions -a fact acknowledged by researchers (Chapelle, 2001a; Warschauer & Healey, 1998), which calls for further investigation.

**Developing CALL Tasks**

Developing activities that would facilitate second language learning has been the concern of researchers since the advent of computers into education. Overcoming the initial scepticism, the educational community shifted from wondering whether we *should* use the computer in second language teaching to *how* we should best use it (Chapelle, 2001a:1). Characteristics particular to CALL tasks have been following parallel routes with the orientation of the educational trends as they developed throughout the years. The early CALL tasks, presented, reinforced and tested
particular language items (Jones & Fortescue, 1987) through simple, computerized versions of traditional question-and-answer, matching or multiple-choice exercises. The limited capabilities of the early computer systems allowed for crude sketches of faces in black and white or combinations of two -the most- colours (see the programs “Photofit” and “Can you guess?” cited in Jones & Fortescue, 1987). Those tasks were grammar-oriented and worked mostly as a motivational supplement to the traditional lesson, assigning the role of knower-of-the-right-answer to the computer (Jones & Fortescue, 1987:5). As computer technology advanced and computers became faster and more powerful, CALL programs became more sophisticated and allowed for a wider range of activities, and a richer graphical environment on the screen. This allowed tasks consistent with the principles of the ‘communicative CALL’ theory as it was introduced by Underwood (1984). He suggested using techniques from artificial intelligence to make the computer recognize learner’s input and generate responses in order to create ‘meaningful’ conversation between computer and learner. The introduction of this technique led to the development of simulations and adventure games, which gave the computer the role of the stimulus as it was defined by Jones & Fortescue (1984:6).

Parallel to the communicative CALL approach, researchers developed the form-focused instruction theory (Long, 1988) and CALL was quickly linked to it. Doughty (1987) set the theoretical principles of taking advantage of CALL to focus on form. She carried out a study comparing “…the effects of explicitly salient L2 input with input which was not explicitly flagged to direct learners’ attention” (Chapelle, 2001a:19). Her findings –“learners receiving salient input performed better on grammatical posttests than did the group receiving normal input”- as Chapelle (2001a) puts it, offered hope for the use of technology in the study of second language acquisition.

Computers allow manipulation of elements on screen and therefore, can help teachers devise numerous ways of ‘flagging’ target items (highlighting, underlining, coloring, rule giving, etc) to direct the learners’ attention to them (Long & Robinson, 1998:18). Therefore, a CALL task that focuses on a discreet point should comprise ‘eye-catching’ elements to make learners notice its grammatical or syntactical behaviour and consequently help them understand its structure and function more effectively.
Animation
The newer instructional software developed and used in CALL has become more complex and more sophisticated, and as a result, teachers are equipped with more choices when it comes to developing CALL tasks. Features such as highlighting, underlining, or coloring that Long & Robinson suggested using as ‘flagging’ vehicles have been enriched with the addition of newer ones, more striking to the eyes of the students. One such feature, which has been found to add to task effectiveness, is animation. Animation with or without text or sound can now play an important role in helping teachers find ways of ‘flagging’ items so that learners can focus on the form that is being taught. Studies (Baek & Layne, 1988; Park & Gittelman, 1992) have verified positive learning effects in certain circumstances where computer animation is used in courseware. Richards & Fukuzawa (1989:28) say, “Well-designed lessons with animation have improved student scores more than presentations with graphics and text or text alone”. Further studies investigating differences in effectiveness between animation with text and animation with sound, have found that the richer the task environment is, the more beneficial to the student it becomes. The findings of an experimental study investigating which type of multimedia is more helpful for learners reveal that “subjects in the animation+text+voice group scored significantly higher on the posttest than those in the animation+voice or the animation+text groups” (Chuang, 2002).

Research Methods
Chapelle sets the principles that should underpin experimental studies which investigate CALL’s effectiveness by saying: “The most convincing way to demonstrate the language learning potential of a CALL activity is through the study of learning outcomes” (2001a:74). She suggests that researchers should conduct a pretest ensuring that the learners who will be involved in the CALL activity -the “treatment group”- did not know the target forms before beginning to work with CALL. Then, upon completion of the task, a posttest should be given to the learners, which would provide evidence for the language learning potential of the task. In order to make the evidence even stronger, the researcher should employ a contrasting group -the “control group”- who will not use the CALL task at all, or use the task in another form, and then compare the differences in gains (Chapelle, 2001a). This method of conducting a study is acknowledged as the most commonly used experimental design although the potential nonequivalence between the groups complicates its statistical analysis (Trochim, 2002). The referral point for the pre and posttests must be a discrete variable, as it is measured objectively and therefore, can help the researcher establish relationships and causality (Egbert & Jessup, 1996).
Conclusion
It is evident that the introduction of computer technology into the language teaching processes armed educators with a powerful weapon to assist them in their struggle for more effective education.

The various studies on CALL’s effectiveness have mostly focused on macro-evaluating learning processes and outcomes. However, although it has been found that computers can potentially enhance isolated linguistic input, this area of CALL has been overlooked. A need to micro-evaluate CALL arises, since advances in computer technology allow educators to easily customize existing software or even develop new one, so that it can focus on isolated forms.

The same research methods that have been employed in educational studies so far, can be employed to micro-evaluate CALL’s effectiveness when it comes to pre-planned itemized focus. These procedures are followed and described in the following chapters.
Chapter III - Methodology
Experiment Design

Trochim (2002) says that “…much contemporary social research is devoted to examining whether a program, treatment, or manipulation causes some outcome or result”. He calls this kind of social research ‘a research into a causal relationship’ and he adds that if researchers want to investigate a causal relationship the strongest design with respect to internal validity is the experimental design. This is the research design that we adopted in this study, as we investigate possible differences in learning outcomes between two different teaching methods –CALL and traditional instruction. The hypothetical causal relationship between the teaching methods and learning outcomes is investigated employing a pretest-posttest randomized experimental design usually referred to as ANCOVA (analysis of covariance) design (Trochim, 2002).

The steps that are followed in this study are the ones that Trochim (2002) describes in ‘The Research Methods Knowledge Base’. Specifically,

- The population is first defined and described.
- The sample drawn from the population is then randomly selected.
- The participants are randomly assigned into a control and a treatment group.
- A pretest is given to both groups.
- The two groups are given the same linguistic input in two different ways.
- A posttest is given to check the effectiveness of the teaching methods.

Population

This study took place in a Greek State Lykeio in the area of Peristeri, which is a large district in Athens, maintaining the characteristics of a typical Greek urban area. The study compares the learning outcomes from two groups of randomly selected Lykeio students and as it was conducted solely by the author, administrative limitations and time restrictions confined the accessible population to the students in the second and third years of the school.

Lykeio comprises a three-year cycle of studies leading to the university. It follows the 9-year obligatory study cycle for the Greek educational system and its students’ age spans from 15 to 17. The students attend English lessons in the last three years of primary and the first three years of secondary education totaling six years of English classes within the framework of public education before coming to Lykeio, although a large number also attend extra-curricular lessons -throughout these years- in private institutions. This leads to lack of homogeneity in the knowledge level of English among Lykeio students, a characteristic acknowledged by the Greek Ministry of Education (Hellenic Republic Government Gazette Issue Two, 1999).
Selecting a Homogeneous Group

The sampling frame for the study was the 103 students in the second and third classes of the school. Before sampling the population, we needed to study the distribution of the students’ knowledge level of English at the time of the experiment. Thus, Oxford Placement Test - a generally accepted as a reliable standardized placement test (Grundy, 1998; Morell Moll, 1999) - was given to all the students. The test consisted of 200 items mostly focusing on grammar and vocabulary. For various reasons 20 students were not able to sit the placement test, therefore response was obtained and analyzed from 83 subjects. The scores ranged from 40% to 88%, with a mean of 68.03% (TABLE 1).

<table>
<thead>
<tr>
<th>TABLE 1: GENERAL DESCRIPTIVE STATISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCORE</td>
</tr>
<tr>
<td>Number of subjects</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>

The findings are consistent with the acknowledged by the Ministry of Education typical picture of an English class in a state Lykeio as there is a quite extensive grade distribution within the 47 ‘B’ class and the 36 ‘C’ class students. This distribution is almost equal for both classes and the average grade for C class is higher than the average grade for B class. Specifically, a range of 45 points and an average of 63.12% was noted for B (second) class, while a range of 42 points and an average of 74.44% was observed for C (third) class (TABLE 2).
TABLE 2:
CLASS DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>CLASS</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error of Mean</th>
<th>Distribution of score</th>
</tr>
</thead>
<tbody>
<tr>
<td>B (N=47)</td>
<td>63.12</td>
<td>12.33</td>
<td>1.79</td>
<td>45.00</td>
</tr>
<tr>
<td>C (N=36)</td>
<td>74.44</td>
<td>9.07</td>
<td>1.51</td>
<td>42.00</td>
</tr>
<tr>
<td>Total (N=83)</td>
<td>68.03</td>
<td>12.34</td>
<td>1.35</td>
<td>48.00</td>
</tr>
</tbody>
</table>

Since score distribution for both classes is almost equal, it was decided to handle the data from both classes as one group, representative of the average knowledge level of English for B and C class Lykeio students.

Trochim (2002) says that “if you have a very diverse group of respondents, they are likely to vary more widely on your measures or observations”. In order to limit diversity we opted to conduct the experiment with a more or less homogeneous group of students. Thus, a level close to the overall mean was arbitrarily selected. The overall average score (68.03–TABLE 2) was set as a starting point and after applying the 65-95-99% rule (Trochim, 2002) it was found that scores from 66.68% to 72.04% account for the 99% of the accessible population. As the number of students who fell within this range was quite small, it was decided to extend the range towards the two extremes so that more students would be employed. Therefore, the list was extended to cover the span from 53% to 79%, which meant that scores below 50% and over 80% were excluded as extreme cases representative of very small proportions of the population.

The final shortlist statistics table (TABLE 3) shows that the list included 54 students in a quite homogeneous group, as their Oxford Placement Test scores ranged from 53% to 79%. This gives a mean of 68.5% (very close to the overall mean in TABLE 1) and a score distribution of 26 points (much lower than that of the whole group in TABLE 1 or the individual classes’ in TABLE 2).
Therefore, by using Oxford Placement Test, we managed to establish the level of knowledge of English of the students who comprised the accessible population for our experiment. We then excluded the students falling in the area below 50% and above 80% and we managed to come up with a quite homogeneous group from which we were going to randomly sample the students who would take part in the experiment.

**Instrumentation**

As mentioned in chapters I and II, the study investigates the effectiveness of using technology to teach isolated linguistic items (forms). The form selected to be used in the experiment was ‘inversion’ -as it is included in the curriculum for the Greek Lykeio.

In order to investigate prior knowledge of the specific form, the students in the shortlist were given a pretest (Appendix I). The pretest was first examined for content validity by two experts in the field of testing (Appendix II) and a few adjustments were made based on the experts’ recommendations. The final form of the pretest comprised 5 different types of exercises and a total of 42 items, each scored objectively. As we did not want to draw the students’ attention to inversion but rather to make them see it as a ‘general grammar test’, we chose to design the test to include various forms (e.g. passive voice, relatives, conditionals etc.). Therefore, from the 42 items only 24 -scattered in various exercises- checked knowledge of inversion and those were the ones that were actually checked, neglecting the items that focused on the other forms.
**Piloting and Pretesting**

Before giving the pretest to the students in the final shortlist, we piloted it with a different group of students in another Lykeio. The piloting did not reveal any weaknesses in the wording of the rubrics or the structure of the exercises and the test seemed clear to the students. The time they required to complete the test was measured and it coincided with our estimate (about 35 minutes on average). Since the piloting did not lead us to any further alterations, we decided to proceed with the next phase of the experiment and give the pretest to the students in the final shortlist.

The results of the pretest showed a very extensive distribution of grades (TABLE 4), which meant that there were students who had already mastered the form, students who were not so confident, and students who totally lacked knowledge of the form. Scores of 0% on one extreme with scores of 79.16% on the other gave a mean at 28.6% and a distribution of 79.16 points.

**TABLE 4:**

<table>
<thead>
<tr>
<th>DESCRIPTIVE STATISTICS – OVERALL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRETEST</strong></td>
</tr>
<tr>
<td>Number of students</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>

**Random Selection**

Based on the results of the pretest, the 40 students with the lowest scores ranging from 0% to 37.5%, were selected to compile the final list from which the students who would participate in the experiment would be randomly selected. The resultant group represented a homogeneous class -in terms of level of knowledge of English- with minimal knowledge of ‘inversion’. This is evident from the data analysis as seen in TABLE 5. We can see that the mean for the placement score for those students was 67.50%, which was very close to the mean of the whole group (68.03% in TABLE 1) and the mean for the pretest checking knowledge of inversion was 19.26% (TABLE 5).
The experiment was to be conducted in a traditional classroom for the first group and in the school’s computer lab for the second. The lab houses 12 students, thus 24 students were needed for the experiment altogether (12 for the lab and another 12 for the classroom). Using Excel’s RAND command (Excel’s way of assigning random numbers to items on a list), each student on the final list of 40 was given a random number. The list of the students was then rearranged from the lowest to the highest random number that Excel assigned to each subject. This method of assigning random numbers to listed subjects and then conducting a random sampling procedure is acknowledged by Trochim, (2002) as the easiest but still valid way to carry out a simple random sampling procedure.

The first 12 boys and the first 12 girls of the rearranged list were then selected to form the two groups (control and treatment) that were needed for the experiment. The remaining students were going to be used as a pool for substitutes.

**Random Assignment**

After the sample of the 24 randomly selected students was set, Excel was once again employed to assign random numbers to the students. The same procedure with random selection was followed and the list was rearranged from the lowest to the highest random number that Excel assigned to each student. The first 6 boys and the first 6 girls that we met as we read down the list were assigned to form group one (the traditional group) and the remaining 12 were listed as group two (lab group). Thus, the 24 students were randomly assigned to two different groups. This random assignment procedure (randomly assigning participants to treatments) is fundamental to any experimental design (Trochim, 2002).

<table>
<thead>
<tr>
<th></th>
<th>PLACEMENT</th>
<th>PRETEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Range</td>
<td>26.00</td>
<td>37.50</td>
</tr>
<tr>
<td>Minimum</td>
<td>53.00</td>
<td>.00</td>
</tr>
<tr>
<td>Maximum</td>
<td>79.00</td>
<td>37.50</td>
</tr>
<tr>
<td>Mean</td>
<td>67.50</td>
<td>19.26</td>
</tr>
</tbody>
</table>
The whole procedure (giving the pretest, correcting it, randomly selecting subjects and randomly assigning groups) was completed in less than two days, so that chances of students’ doing a lesson on inversion outside the Lykeio framework would be minimized.

Software Design
The software (Appendix III) used for the presentation of the lesson, was prepared by the author. It comprised a series of webpages, which were compiled into a website and published on the Internet (http://users.forthnet.gr/ath/vasdor/ma/start.html). The webpages presented the form ‘inversion’ in an interactive way, where the students would be able to explore -employing ‘point and click’ techniques- the various cases where inversion can be used.

Two grammar books (Quirk & Greenbaum, 1973; Evans, 1998) were consulted so that ‘inversion’ cases would be listed and presented with examples. The six most common cases were selected to be presented to the students, and the example sentences that were taken from the grammar books were listed both in inverted and non-inverted form so that the students would get a clear picture of the mechanisms involved in inverting sentence elements for emphasis. The grammar rules for inversion together with the examples made up the ‘inversion website’ that the author published on the Internet. Taking advantage of the animation features and multimedia capabilities of the software packages ‘Macromedia Fireworks’ and ‘Macromedia Flash’, the changes in the syntax of the sentences when they are inverted were ‘flagged’ (as it was explained in chapter II) to draw students’ attention on them.

After a brief introductory page (Appendix IVa), a basic definition of ‘inversion’ is given (Appendix IVb) and the users are led to the starting page of the website (http://users.forthnet.gr/ath/vasor/ma/navbar.html) (Appendix V). This page was designed to constitute the backbone of the site. It employs a dynamic menu that leads to the basic ‘chapters/sections’ of the whole website.

The rules of inversion together with examples for each case are found in the ‘Inversion cases’ section. The six selected inversion cases are presented in separate web pages, which can be accessed from the starting page. In each case, the same ‘flagging’ techniques were employed; the key elements of the form were stressed by different colour typeface and the changes were shown through animation (Appendix VI a & b).

The ‘Examples’ section follows, where authentic examples of inversion cases taken from a song (‘No More’ by Robertson & Blair sung by Elvis Presley) are presented, employing a combination of animation and sound. This part was designed to solidify the students’ perception of the form,
leaving them with a vivid impression of the example sentences due to the combination of animation and the familiar sound of the song.

The last section of the website is the ‘Practice’ section (Appendix VII). A multiple-choice and a matching exercise checking knowledge of the form ‘inversion’ had already been prepared by the author. Using the ‘Hot Potatoes’ package, they were compiled in webpage format and they were incorporated into the basic ‘inversion’ website. The ‘hot potato’ exercises were meant to be self-assessment instruments, as they provide feedback on the students’ performance.

The lessons
The two different lessons on ‘inversion’ were designed and taught by the same teacher -the author- eliminating the need to take into consideration variables that revolve around the teacher. Long & Javidi (2001) reported that variables such as instructor expertise in the course content, clearly communicating course content, making complex theory easier to understand, or effectively using technology required further adjustments in the analyses of the results of their study.

Three days after the pretest, the first group of students (the lab group) was led to the computer lab (Appendix VIII) where the students were first given instructions on how to navigate around the web pages. Each student was seated in front of a computer screen so that they could work independently. This seating arrangement has been criticized (Jones & Fortescue, 1987) as counter-productive because it isolates students and does not promote collaboration; however, it was preferred because we were looking into the effectiveness of technology on individual subjects.

The students were given instructions to navigate the website following the links in the order given in the starting page. Therefore, after they exploited the ‘Inversion Cases’ section, they moved on to the ‘Examples’ section and they ended up exploiting the ‘Practice’ section. The students were left to spend as much time –within the timetable restrictions- as they felt they needed going back and forth through the various sections of the website. They revisited the ‘examples’ section to make sure they had understood the syntactical mechanism pertaining to inverting sentence elements or they repeated the ‘practice’ section until they saw they could answer all the questions correctly. The whole procedure lasted for a full teaching period of 45 minutes.

The second group (the traditional group) of students was scheduled for the following teaching period. The students were assembled in a regular classroom instead of the computer lab, and the
lesson was delivered based on exactly the same linguistic input. The example sentences presenting the inversion cases and the multiple choice and matching exercises in the practice section were the same as the ones used for the ‘lab’ group. All the material was presented on the blackboard and the teacher acted as a ‘knower-of-the-right-answer’, a role that Jones & Fortescue (1987:5) had saved for the computer. Questions from individual students were answered after the completion of all the sections, when the whole lesson was reviewed in brief. This procedure took the same time as the procedure in the lab; one teaching period of 45 minutes.

**The Posttest**

The posttest was conducted two days after the lessons. We opted for an immediate evaluation of the lessons’ effectiveness because we were not interested in the retention factor. The posttest (Appendix IX) was the same test that was given as pretest, the only difference being the fact that all ‘unnecessary’ items (items not checking knowledge of ‘inversion’) were taken out. The choice of focusing on inversion only, was made after considering the fact that the students would immediately know it was a test on inversion since the material was presented only two days before. Thus, the test was a pure discrete point test, focusing only on the form ‘inversion’.

**Data Analysis**

The data retrieved from the placement test, the pretest and the posttest were analyzed in the statistical software package ‘SPSS’. The following analyses were conducted:

- A T-Test investigating overall differences between the pretest and the posttest findings.
- An ANCOVA investigating whether there were differences in the outcomes of the two teaching methods, and
- An ANOVA investigating whether there were differences in the performance between males and females.
Chapter IV – Presentation and Analysis of Data
The retrieved data were analyzed in the statistical software package ‘SPSS for Windows’. The scores (placement, pretest and posttest) were treated as numeric variables under their own names. The two groups representing the two different teaching methods were given the values ‘0’ -for the traditional or control group- and ‘1’ -for the lab or treatment group; thus, a numeric variable was created. It is referred to here as a ‘dummy variable’ since it artificially uses a number to represent categories (Carver & Nash, 2000). The same procedure was followed for the discrimination between males and females. A dummy variable was created with values of ‘0’ and ‘1’ for female and male students respectively.

Descriptive Statistics
The control and treatment groups consisted of 12 students each, randomly selected from the pool of the 40 students who scored lowest on the inversion test (the pretest). When one looks at the results for all 24 students (TABLE 6), one notices a significant increase in students’ performance, which rose from a mean score of 17.70 in the pretest to a mean score of 64.23 in the posttest. A wider grade dispersion in the posttest scores (range 79.16 vs. 33.33) is also noticed, but this is considered normal, as the scores in the posttest are higher and according to Carver & Nash (2000) whenever an analyst obtains posttest scores higher than the ones in the pretest, the dispersion for the posttest is usually wider.

TABLE 6:
PRETEST, POSTTEST DESCRIPTIVE STATISTICS FOR ALL STUDENTS

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Range</td>
<td>33.33</td>
<td>79.16</td>
</tr>
<tr>
<td>Mean</td>
<td>17.70</td>
<td>64.23</td>
</tr>
</tbody>
</table>

An analysis of frequencies for the same findings (TABLE 7) shows that the median and the mode for the pre and posttests follow the same trend and confirm the increase in the overall performance of the students.
TABLE 7:
MEAN, MEDIAN AND MODE FOR ALL STUDENTS

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>17.70</td>
<td>64.23</td>
</tr>
<tr>
<td>Median</td>
<td>16.66</td>
<td>70.83</td>
</tr>
<tr>
<td>Mode</td>
<td>12.50</td>
<td>83.33</td>
</tr>
</tbody>
</table>

The T-test
The first thing that was investigated was whether the differences between the pre- and posttest for both groups were significant. The appropriate statistical test for instances involving “the change in a single variable observed at two points in time” is a paired samples t-test (Carver & Nash, 2000:128).

In order to conduct the T-Test, a null hypothesis was formulated, which stated that ‘there were no significant differences between the pre- and the posttest scores’. This null hypothesis was checked against the alternative hypothesis that ‘the differences between the pre- and posttest scores were significant’.

For this t-test, as well as throughout the whole study, the P-value (significance level) was set at $\alpha=0.05$ (95% confidence), which is a generally accepted standard for all statistical analyses and a “rule of thumb for all social research” (Trochim, 2002).
The t-test analysis (TABLE 8) showed that the difference in the means between the posttest (M= 64.23, SD= 21.09) and the pretest (M=17.7, SD=8.96) is statistically significant (t= -9.28, df= 23, 2-tailed p<.05) because the P value (.000) is lower than .05. Therefore, the null hypothesis is rejected, the difference is statistically significant, and as the possibility of attributing it to chance is ruled out, it is attributable to the two teaching methods.

The question remaining to be answered is whether there is a difference in the separate means for the control and treatment groups; in other words, if one of the two teaching methods yielded better results than the other. The posttest descriptive statistics (TABLE 9) show that the overall mean increased from 17.70 to 64.23; however, a comparison of the means for the two groups shows that the control group (the ‘traditional’ group) scored higher in the posttest than the treatment group (the ‘lab’ group). Specifically, the mean for the control group was 73.61, whereas the mean for the treatment group was 54.86. This might imply that the traditional method of teaching was more effective, but in order to check its statistical significance, further analysis was required.
A visualization of the results for the pre- and posttests for both groups is presented in Figure 1. The improvement in performance from the pretest to the posttest for both groups is striking. The fact that the control (traditional) group scored higher than the treatment (computer lab) group in the posttest, is also obvious. In order to see if the difference between the two groups is significant and attributable to the different teaching methods, additional statistical analyses were conducted.

FIGURE 1: PRETEST, POSTTEST MEANS OF THE GROUPS

ANCOVA
In our study differences between the pre- and posttests were recorded, so it was decided to employ an analysis of covariance (ANCOVA) to check if those differences are significant. Krus (2002) says that “differences in achievement scores are attributable not only to differences
among the teaching methods but also to the initial differences in pretest scores. In order to control this source of variability (the covariate variable), the statistical technique - analysis of covariance (ANCOVA) is used. Analysis of covariance adjusts achievement scores based on aptitude scores (the covariate)”.

There are five variables analyzed in the study. Three numeric ones (the placement test, the pretest and the posttest scores) and two dummy variables representing categories (the teaching method and the sex). The former are the dependent variables of the study -not manipulated by the researcher- and the latter are the independent ones (Trochim, 2002). One more numeric variable was created after subtracting the pretest from the posttest. This variable showed the difference in the performance of the students and it was assigned the name ‘gain’. The ANCOVA checked the significance of the dependent variable ‘gain’, after it was controlled for the pretest covariate.

We decided to follow Lane’s (2002) steps in hypothesis testing; therefore, the null hypothesis was first specified: ‘There is no difference between the two teaching methods’ (H0: μ1 - μ2 = 0). The alternative hypothesis is: ‘The two methods yield different results’ (H1: μ1 ≠ μ2). The next step was to select a significance level. As it has already been stated, the significance level for all the analyses in this study was set at α = 0.05.

### TABLE 10:
GAIN DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>Teaching Method</th>
<th>N</th>
<th>Mean for ‘gain’</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Lesson</td>
<td>12</td>
<td>13.67</td>
<td>4.83</td>
</tr>
<tr>
<td>Computer Lab</td>
<td>12</td>
<td>8.67</td>
<td>5.97</td>
</tr>
</tbody>
</table>
The results of the ANCOVA indicated that there were significant differences between the two teaching methods. The descriptive statistics in TABLE 10 show that the control group had a higher gain than the treatment group. The difference between the group means after covarying the pretest scores, is significant, $F= 5.09, p= .035$ (TABLE 11). Thus, the null hypothesis is rejected in favour of the alternative hypothesis that the two methods yielded different results.

The ANCOVA allows us to say that there is a 95% probability that the gain is not attributed to chance. Since TABLE 10 shows that the ‘traditional’ group had a higher gain, we can presume that there is a 95% probability that the ‘traditional’ method of teaching the form ‘inversion’ to students in a Greek State Lykeio, is more effective than teaching it in a computer lab.

**Differences between males and females**

The subsidiary research question of this study is whether there are any significant differences in the performance between males and females. The gain means for the two groups (TABLE 12) show that there is a slightly higher gain in favour of the male group (11.68 vs. 10.86).
TABLE 12:
DESCRIPTIVE STATISTICS FOR THE GAIN MEANS BETWEEN MALES AND FEMALES

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>10.86</td>
</tr>
<tr>
<td>Males</td>
<td>11.68</td>
</tr>
</tbody>
</table>

In order to check if this difference was significant, an ANOVA for the variable ‘sex’ was employed. The null hypothesis is that there were no significant differences between males and females and the alternative hypothesis is that the two sexes performed differently. The significance level was set to $\alpha = .05$.

TABLE 13:
ANOVA FOR ‘SEX’

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.788</td>
<td>1</td>
<td>.78</td>
<td>.022</td>
<td>.884</td>
</tr>
<tr>
<td>Within Groups</td>
<td>798.54</td>
<td>22</td>
<td>36.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>799.33</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results (TABLE 13) show that the p value is greater than our $\alpha$ level ($F = .022, p = .884$) which means that the null hypothesis is not rejected and the differences noted between males and females are not significant.

A separate analysis was conducted to investigate whether there were significant differences in the performance of males and females in each separate group, in other words whether males or females responded differently in the different teaching methods. In order to carry out this analysis, a two-factor independent measures ANOVA was employed. The null hypothesis is that there are no differences in performance between males and females for each group.
The ANOVA results (TABLE 14) show that the p value for the variables sex and teaching method is greater than .05 \((F=.622, p=.440)\). Therefore, we cannot reject the null hypothesis and we can say that there is a 95\% probability that the differences noted for males and females in the two teaching methods are attributed to chance.

Recapitulating the analyses results, we can say that there were significant differences between the two teaching methods, favouring the traditional group. The differences noted between males and females for either teaching method are not significant.
Chapter V – Conclusions and Recommendations
Conclusions
This study’s primary aim was to investigate whether computer-mediated instruction was more, less than, or as effective as traditional teaching in a Greek State Lykeio. Our prediction was that implementing computer-mediated instruction would yield different results from implementing traditional teaching methods.

The null hypothesis we formulated in order to investigate this prediction, was that the learning outcomes of the two different teaching methods would be the same. The alternative hypothesis was that one of the two teaching methods would be more effective. Since we did not specify a direction in our prediction –in other words, our prediction was not in favour of either of the methods- we came up with a two-tailed hypothesis; we hypothesized that computer-mediated instruction would be either more or less effective than traditional teaching.

The results analyzed in chapter IV led us to reject the null hypothesis in favour of the alternative one and conclude that the two teaching methods give different outcomes. A further analysis of the learning outcomes showed a significant difference favouring the traditional method of teaching.

A subsidiary question that this study investigated was if computer-mediated instruction was equally effective for males and females. This question was answered on the basis of the null hypothesis that computer-mediated instruction presented no different learning outcomes from traditional teaching. The alternative hypothesis was that the two methods produced different outcomes. The analysis of the findings in chapter IV led to the acceptance of the null hypothesis and to the conclusion that CALL and traditional teaching present no significantly different learning outcomes between males and females.

Inferences
In his analysis of social research methods, Trochim (2002) states that most of the inferential statistics come from the general family of statistical models known as ‘the General Linear Model’ (GLM). The statistical analyses that we conducted in this research –namely the T-test, the ANOVA and the ANCOVA- belong to this family of statistical models. They were selected because we wanted to generalize the results to the population from which the sample was drawn. Thus, we can infer that a student in a Greek State Lykeio would benefit more from traditional methods of teaching -as opposed to CALL, when learning isolated formS. Additionally, we inferred that there are no differences for males and females in the effectiveness of the two teaching methods.
However, these inferences encompass some fine points that require further clarification. The administrative limitations during the experiment confined us to a small sample in a specific area of Athens. We cannot assume that the same experiment in different areas -with different population characteristics- would yield the same results. Furthermore, the two ‘classes’ we formed after the random sampling and assignment were homogeneous. Specifically, they were made up of students whose level of knowledge of English lay anywhere within the span 50% to 80% as measured in Oxford Placement Test, which is not the case in the majority of the Greek State Lykeia. Therefore, we cannot be certain if a heterogeneous class would respond in the same way in the two different teaching methods.

Time limitations restricted the study on the effectiveness of teaching ‘inversion’ only. This form is used mainly for emphasis (Evans, 1998) and it is not loaded with communicative value; communication between two interlocutors can be accomplished without using inversion at all. Thus, teaching forms that would be essential for communication might bring different outcomes.

Implications of the Study
All students improved significantly due to the two teaching methods. The students in the ‘traditional’ group did significantly better than the students in the ‘lab’ group. Since both groups worked equally enthusiastically, we were led to attribute the difference to the fact that the Greek students are still quite attached to ‘traditional methods’ when it comes to teaching. Informal unrecorded discussions with the students in the two groups gave us the impression that the ‘lab’ group enjoyed working on the computers and found the lesson ‘very interesting’. On the other hand, the traditional group was more focused on the lesson because the students saw it as part of their everyday routine, which does not include working in the computer lab.

CALL did benefit students. However, the study led us to believe that computer lab activities need to be supported with ‘traditional’ elements of teaching -like for example explanations on the board or attention to individual questions. The differentiating characteristic of the lab group was that the students worked on their own and tried to grasp the particulars of inversion by going back and forth to the web pages presenting the rules and the examples. The traditional group on the other hand, had the chance to exchange ideas and listen to answers as a group rather than individually. This, probably helped students to clarify more fine points than they had originally thought of, as they could listen to their peers’ remarks.
Further Research and Recommendations

As noted earlier, in order to come to more definitive conclusions, a more elaborate study is required. Larger samples from various areas of Greece should be involved so that different characteristics of the societies in various areas could be taken into account. The research should also involve randomly selected heterogeneous groups, without prior control for knowledge level. Furthermore, it should investigate a number of forms—including grammar, syntax, or vocabulary, with linguistic input of more communicative value. Finally, a factor that has not been touched in this study—due to time limitations—but we think is quite important when investigating learning outcomes, is retention.

Recapitulating, our opinion is that a more elaborate study on CALL effectiveness should involve more geographical areas, larger and heterogeneous groups, presentation of a larger number of forms and investigation of retention.
Bibliography


Nesler, M. S. & Hanner M. B. (2001) *Professional Socialization of Baccalaureate Nursing Students: Can Students in Distance Nursing Programs Become.* Unpublished. Accepted for 2001 publication in the Journal of Nursing Education


Appendices
Appendix I – The Pretest

Name:.............................................................................................................................................

A. Choose the correct alternative in order to complete the sentence.

1. Football ______________ by a lot of men.
   a. enjoys
   b. is enjoying
   c. is enjoyed
   d. enjoyed

2. Not until the last _______________ able to relax.
   a. guest left, we were
   b. guest had left, we were
   c. guest had left, were we
   d. guest left, did we

3. This painting _______________ to be his masterpiece.
   a. believes
   b. is believed
   c. believed
   d. did believe

4. Only if you agree _______________ with the plan.
   a. can we go on
   b. we can go on
   c. we will go on
   d. we must go on

5. Rarely _______________ to the cinema.
   a. I go
   b. I do not go
   c. do I go
   d. I am going

6. Not since the 30s _______________ so much snow.
   a. there is
   b. there has been
   c. has there been
   d. there was

7. If you had come earlier, _______________ seen Mary.
a. you could be
b. you might be
c. you will have
d. you could have

8. No sooner __________________________ than the sun came out.
   a. had it stopped raining
   b. it stopped raining
   c. it didn’t stop raining
   d. it had stopped raining

9. The students ______________ new assignments by the teacher.
   a. will give
   b. are giving
   c. have given
   d. have been given

10. At no time ______________ of anything unusual.
    a. I was aware
    b. was I aware
    c. I am aware
    d. I will be aware

B. Put the words in the correct order to produce complete sentences beginning with the bold underlined word/phrase.
   a. sun here comes the
   b. had I hair week done last my
   c. she never will lies tell
   d. working he be believed to illegally is
   e. leave should early you tell could me? you
   f. they finished only after did painting house move into their they had new
   g. didn’t say why you so?
   h. others so I worked did hard the and
   i. you what did to tell he do?
   j. by only hard the can you studying pass exam.

C. Choose whether the phrase is right or wrong by ticking the correct box
<table>
<thead>
<tr>
<th></th>
<th>Right</th>
<th>Wrong</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Here my father comes for work.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Never before have I seen such bad weather.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>If you had been more careful, you wouldn’t damage the car.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Not once has he sent me a postcard.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>So well she played, that everyone was amazed.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Only if you help me can we move this piano.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Fast though did they ran, they couldn’t catch us.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>There goes the bus.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>She is more ambitious than her brother.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Only if you come here can we discuss the plan.</td>
<td></td>
</tr>
</tbody>
</table>

D. Rewrite the following sentences so that they mean the same.

a. I didn’t remember to call Alexandra until it was too late.
   Not until ____________________________________________

b. He could only reach the shelf by standing on a chair.
   Only by _____________________________________________

c. He is a fine tennis player and a good footballer.
   Not only ____________________________________________

d. I’ve never had such fun anywhere else.
   Nowhere ____________________________________________

e. He knew little about the truth.
   Little ______________________________________________
KEY

A

<table>
<thead>
<tr>
<th>1.c</th>
<th>2.c</th>
<th>3.b</th>
<th>4.a</th>
<th>5.c</th>
<th>6.c</th>
<th>7.d</th>
<th>8.a</th>
<th>9.d</th>
<th>10.b</th>
</tr>
</thead>
</table>

B.

a. Here comes the sun.
b. I had my hair done last week
c. Never will she tell lies
d. He is believed to be working illegally
e. Should you leave early could you tell me?
f. Only after they had finished painting did they move into their new house.
g. Why didn’t you say so?
h. I worked hard and so did the others.
i. What did he tell you to do?
j. Only by studying hard can you pass the exam.

C.

<table>
<thead>
<tr>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
</tr>
</thead>
<tbody>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

D.

1. Not until it was too late did I remember to call Alexandra.
2. Only by standing on a chair could he reach the shelf.
3. Not only is he a fine tennis player, but also a good footballer.
4. Nowhere else have I had such fun.
5. Little did he know about the truth.
Appendix II - Pretest content validity judges

- Dr. A. Taglides  
  EFL State School Advisor,  
  Tutor at the Hellenic Open University

- Dr. C. Agiakli  
  Tutor at the Hellenic Open University
Appendix III – The Software Used
Appendix IV – Introductory pages

This presentation is about a grammatical phenomenon which is called...

INVERSION

INVERSION

Sometimes when we want to place emphasis on a phrase, we INVERT (change the order of the subject and verb) and we make the phrase LOOK LIKE a question, although it isn’t.

This grammatical phenomenon is called INVERSION.

Let’s see when we can use it.
Appendix V – The Starting page
Appendix VI – Rule Presentation

You invert subject and verb...

after time and manner expressions (e.g. only after..., only by..., only when...)

Example:

You can go out only after you finish.

GO BACK START

You invert subject and verb...

after time and manner expressions (e.g. only after..., only by..., only when...)

Example:

only after you finish can go out

GO BACK START
Appendix VII – Practice Section

**INVERSION: Choose the correct inverted alternative.**
Multiple-choice exercise

1. If I had known about it, I wouldn't have come.

A. I had known about it, I wouldn't have come.
B. Had I known about it, I wouldn't have come.
C. Had I known about it, I wouldn't have come.

Next Exercise
Appendix VIII – The Computer Lab
Appendix IX – The posttest

A. Choose the correct alternative in order to complete the sentence.

1. Not until the last _____________________ able to relax.
   a. guest left, we were
   b. guest had left, we were
   c. guest had left, were we
   d. guest left, did we

2. Only if you agree ________________ with the plan.
   a. can we go on
   b. we can go on
   c. we will go on
   d. we must go on

3. Rarely ________________ to the cinema.
   a. I go
   b. I do not go
   c. do I go
   d. I am going

4. Not since the 30s ________________ so much snow.
   a. there is
   b. there has been
   c. has there been
   d. there was

5. No sooner ________________ than the sun came out.
   a. had it stopped raining
   b. it stopped raining
   c. it didn’t stop raining
   d. it had stopped raining

6. At no time ________________ of anything unusual.
   a. I was aware
   b. was I aware
   c. I am aware
   d. I will be aware
B. Put the words in the correct order to produce complete sentences beginning with the bold underlined word/phrase.
   a. sun **here** comes the
   b. she **never** will lies tell
   c. leave **should** early you tell could me? you
   d. they finished **only after** did painting house move into their they had new
   e. others so **I** worked did hard the and
   f. by **only** hard the can you studying pass exam.

C. Choose whether the phrase is right or wrong by ticking the correct box

<table>
<thead>
<tr>
<th></th>
<th>Right</th>
<th>Wrong</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Here my father comes for work.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Never before have I seen such bad weather.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Not once has he sent me a postcard.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>So well she played, that everyone was amazed.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Only if you help me can we move this piano.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>There goes the bus.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Only if you come here can we discuss the plan.</td>
<td></td>
</tr>
</tbody>
</table>
D. Rewrite the following sentences so that they mean the same.

a. I didn’t remember to call Alexandra until it was too late.
Not until ____________________________________________

b. He could only reach the shelf by standing on a chair.
Only by _____________________________________________

c. He is a fine tennis player and a good footballer.
Not only _____________________________________________

d. I’ve never had such fun anywhere else.
Nowhere ___________________________________________

e. He knew little about the truth.
Little ______________________________________________
KEY

A

| 1.c | 2.a | 3.c | 4.c | 5.a | 6.b |

B.

a. Here comes the sun.
b. Never will she tell lies
c. Should you leave early could you tell me?
d. Only after they had finished painting did they move into their new house.
e. I worked hard and so did the others.
f. Only by studying hard can you pass the exam.

C.

<table>
<thead>
<tr>
<th>1.</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

D.

1. Not until it was too late did I remember to call Alexandra.

2. Only by standing on a chair could he reach the shelf.

3. Not only is he a fine tennis player, but also a good footballer.

4. Nowhere else have I had such fun.

5. Little did he know about the truth.