THE EFFICACY OF CORRECTIVE FEEDBACK
AND TARGET FORM ENHANCEMENT
IN PROMOTING ACQUISITION OF
THE Â/ÂU/EN/AUX DISTINCTION IN L2 FRENCH

by

Paul A. Lyddon

Copyright © 2007 Paul A. Lyddon

A Dissertation Submitted to the Faculty of the

GRADUATE INTERDISCIPLINARY PROGRAM IN
SECOND LANGUAGE ACQUISITION AND TEACHING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2007
THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Paul A. Lyddon entitled *The Efficacy of Corrective Feedback and Target Form Enhancement in Promoting Acquisition of the à/au/en/aux Distinction in L2 French* and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

______________________________  Date: May 15, 2007
Robert Ariew, PhD

______________________________  Date: May 15, 2007
Beatrice Dupuy, PhD

______________________________  Date: May 15, 2007
Joel R. Levin, PhD

______________________________  Date: May 15, 2007
Renate Schulz, PhD

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

______________________________  Date: May 15, 2007
Dissertation Director: Robert Ariew, PhD
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: PAUL A. LYDDON
ACKNOWLEDGMENTS

Though I could never adequately thank everyone I should in such little space, in my meager attempt I would like to begin by duly recognizing my advisor, Dr. Robert Ariew, without whose inestimable mentoring, tireless encouragement, and shared sense of urgency, this project could never have gotten started, much less reached fruition.

I would also like to express my sincere appreciation to the other three members of my dissertation committee—Dr. Beatrice Dupuy, Dr. Joel Levin, and Dr. Renate Schulz—whose specialized professional knowledge and unwavering personal support were likewise essential to my success. Professor Dupuy not only meticulously commented on the numerous drafts of my cloze passages and helped ensure access to my requested pool of research participants, but graciously assumed responsibility for my human subjects recruitment as well. Professor Levin patiently but painstakingly scaffolded my methodological thinking and immeasurably deepened my statistical understanding. Professor Schulz suggested several practical research considerations as well as provided thorough feedback, detailed remarks, and many valuable editing suggestions on the drafts of this manuscript.

Special thanks also go to Dr. Susan Penfield for her direction during my initial feasibility study, to Shaun O’Connor for expediting my human subjects proposal, to Dr. Irène d’Almeida for her authorization to conduct research in the Department of French and Italian, and to Dr. Jerome D’Agostino for his counsel on options for equating alternate test forms.

I wish to thank the following fellow French instructors as well: Delfina De Seixas, for editing and proofreading the texts for my treatment lesson and for beta testing its first functional incarnation; Amandine Andrade, Virginie Dasse-Askildson, Valérie Ferguson, Charlotte Nogier, and Yannick Treguer, for their invaluable input on both my instructional materials and my cloze measures; and Greta Arendsen, Raymond Gnanwo Hounfodji, Emily Mabry, Kathleen O’Brien de Ramirez, Zalia Maiga Touré, and Abir Zihri, for their gracious cooperation and good humor through my many interruptions to their classes.

I also owe an untold debt of gratitude to various members of the College of Humanities Instructional Computing team: Hale Thomas, for the extended loan of a laptop with specialized software; Mark Bryant, for creating my students’ user accounts and profiles; Peter In-Albon, for ensuring smooth network access on the first day of my data collection; and Justin LeBreck, for his Java programming genius and the compilation of my data.

Finally, I wish to thank three very special SLAT classmates—Neil Johnson, Sara Beaudrie, and Hassan Souleyman—for their moral support, their intellectual stimulation and, most of all, their friendship.
DEDICATION

To my darling wife Junko,

whose precious love makes all things possible.
# TABLE OF CONTENTS

LIST OF FIGURES ...........................................................................................................10

LIST OF TABLES .............................................................................................................13

ABSTRACT .......................................................................................................................15

CHAPTER 1: INTRODUCTION.......................................................................................17
   1.1 Historical and Theoretical Background of the Problem.................................18
   1.2 Rationale for the Present Study.........................................................................32
   1.3 Definitions of Important Terms.........................................................................35
   1.4 Outline of the Dissertation................................................................................40

CHAPTER 2: LITERATURE REVIEW............................................................................42
   2.1 The Benefits of Instruction................................................................................42
   2.2 The Need for Negative Evidence........................................................................51
      2.2.1 The Nativist Perspective............................................................................51
      2.2.2 The L2 Debate.........................................................................................57
      2.2.3 Empirical Findings in SLA........................................................................63
         2.2.3.1 General Effects of Form-focused Instruction.................................64
         2.2.3.2 Unique Effects of Corrective Feedback.........................................69
   2.3 The Issue of Explicitness....................................................................................77
      2.3.1 Implicit and Explicit Instruction...............................................................78
      2.3.2 Implicit and Explicit Corrective Feedback.............................................82
   2.4 The Importance of Perceptual Salience...........................................................93
   2.5 Summary.............................................................................................................105
# TABLE OF CONTENTS—continued

## CHAPTER 3: METHODOLOGY

- **3.1 Study Design**
- **3.2 Research Questions**
- **3.3 Hypotheses**
- **3.4 Participants**
- **3.5 Linguistic Target Structures**
- **3.6 Outcome Measures**
  - 3.6.1 General Cloze Procedure
  - 3.6.2 Reading Passage Development
  - 3.6.3 Actual Cloze Construction
- **3.7 Instructional Treatment Materials**
  - 3.7.1 Basic Setup
  - 3.7.2 Feedback Commonalities Across Treatment Groups
  - 3.7.3 Variations Among Treatment Conditions
    - 3.7.3.1 Feedback on *Correct* Short Answers
    - 3.7.3.2 Feedback on Factually Incorrect Answers
    - 3.7.3.3 Feedback on Factually Correct but Grammatically Incorrect Answers
      - 3.7.3.3.1 Versions 1 and 2 (Meaning-focused Feedback, +/- Input Enhancement)
      - 3.7.3.3.2 Versions 3 and 4 (Implicit Form-focused Feedback, +/- Input Enhancement)
      - 3.7.3.3.3 Versions 5 and 6 (Non-metalinguistic Explicit Form-focused Feedback, +/- Input Enhancement)
      - 3.7.3.3.4 Versions 7 and 8 (Metalinguistic Form-focused Feedback, +/- Input Enhancement)
  - 3.7.4 Feedback on Factually Correct but Grammatically Incorrect Answers
- **3.8 Procedures**
TABLE OF CONTENTS—continued

3.9 Data Processing........................................................................................................167
3.10 Summary..................................................................................................................169

CHAPTER 4: RESULTS AND ANALYSES.......................................................................171
4.1 Research Question #1: Learnability of the Linguistic Target Structures.............171
4.2 Research Question #2: Immediate Effects of Corrective Feedback and
   Target Form Enhancement..........................................................................................183
4.3 Research Question #3: Long-term Effects of Corrective Feedback and
   Target Form Enhancement..........................................................................................191
4.4 Summary.....................................................................................................................198

CHAPTER 5: DISCUSSION AND CONCLUSIONS..............................................................200
5.1 Synopsis of Findings.................................................................................................201
5.2 Pedagogical Implications.........................................................................................202
5.3 Limitations of the Present Study.............................................................................211
5.4 Suggestions for Future Research............................................................................217
5.5 Final Word to Fellow Teachers...............................................................................221

APPENDIX A: Project Introduction Script.................................................................223
APPENDIX B: Practice Cloze #1...................................................................................224
APPENDIX C: Practice Cloze #2...................................................................................227
APPENDIX D: Desire to Learn (D2L) Follow-up to Practice Cloze #2.......................229
APPENDIX E: Pretest/Posttest/Delayed Posttest (French-speaking World Version).....232
APPENDIX F: Pretest/Posttest/Delayed Posttest (Muslim World Version)...................236
TABLE OF CONTENTS—continued

APPENDIX G: Unenhanced Treatment Lesson (Integral Version)..........................240
APPENDIX H: Enhanced Treatment Lesson (Story Text Only)..............................313
APPENDIX I: Post-treatment Survey.................................................................327
APPENDIX J: Human Subjects Recruitment Script...............................................336
REFERENCES.................................................................................................337
LIST OF FIGURES

FIGURE 1. Treatment conditions by corrective feedback type and input enhancement level…………………………………………………………………………110

FIGURE 2. List of instructionally presented toponyms and their associated prepositions in chronological order of appearance in the lesson………………131

FIGURE 3. Typical reading passage: Postcard 1 (from France), modeling aux Etats-Unis…………………………………………………………………………………132

FIGURE 4. Short answer question: Part 1 of 3 in follow-up sequence to Postcard 1 (from France), designed to elicit the response aux Etats-Unis ...................134

FIGURE 5. True/false question: Part 2 of 3 in follow-up sequence for Postcard 1 (from France)…………………………………………………………………………..135

FIGURE 6. Multiple-choice question: Part 3 of 3 in follow-up sequence for Postcard 1 (from France)…………………………………………………………………………136

FIGURE 7. Reading passage featuring two target prepositions: Postcard 2 (from the United States), modeling à New York and aux Petites Antilles…………137

FIGURE 8. Lesson versions by corrective feedback type and input enhancement level……………………………………………………………………………………141

FIGURE 9. Short answer question prior to student response…………………………143

FIGURE 10. Feedback to ‘correct’ response on short answer question for all lesson versions (i.e., 1 through 8)…………………………………………………………144

FIGURE 11. Feedback to factually incorrect response on short answer question for odd-numbered lesson versions (i.e., 1, 3, 5, and 7)…………………………146
LIST OF FIGURES—continued

FIGURE 12. Feedback to factually incorrect response on short answer question for even-numbered lesson versions (i.e., 2, 4, 6, and 8)…………………………147

FIGURE 13. Feedback to factually correct but grammatically incorrect response on short answer question for lesson versions 1 and 2 (FoM)………………….149

FIGURE 14. Follow-up providing second model of target preposition for ‘correct’ response on short answer question…………………………………….…..151

FIGURE 15. Enhanced reading passage from lesson version 2: Postcard 1 (from France), modeling aux Etats-Unis in bold red and underlined……………..153

FIGURE 16. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 3 (unenhanced recast)…………154

FIGURE 17. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 4 (enhanced recast)………….155

FIGURE 18. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 5 (unenhanced rejection)…..156

FIGURE 19. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 6 (enhanced rejection)…………..157

FIGURE 20. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 7 (unenhanced rejection with rule explanation)……………………………………………………..……..158

FIGURE 21. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 8 (enhanced rejection with rule explanation)……………………………………………..……………..159

FIGURE 22. Project implementation schedule…………………………………………..161
LIST OF FIGURES—continued

FIGURE 23. Student instructions for treatment lesson.................................165

FIGURE 24. Overall mean target form accuracy scores over time.........................173

FIGURE 25. Immediate posttest target form accuracy by corrective feedback type
by input enhancement level.................................................................189

FIGURE 26. Mean target form accuracy scores by corrective feedback conditions
over time..............................................................................................193

FIGURE 27. Delayed posttest target form accuracy by corrective feedback type
by input enhancement level.................................................................196
LIST OF TABLES

TABLE 1. Student Self-reported Age at End of Course.................................117
TABLE 2. Student Self-reported Reasons for Enrolling in Course ..................117
TABLE 3. Means and Standard Deviations for Target Form Accuracy Scores across All Groups over Time.................................................................172
TABLE 4. Repeated-measures ANOVA Results for Corrective Feedback Type, Input Enhancement, and Testing Order on Pretest and Immediate Posttest Target Form Accuracy.........................................................174
TABLE 5. Repeated-measures ANOVA Results for Corrective Feedback Type, Input Enhancement, and Testing Order on Immediate and Delayed Posttest Target Form Accuracy.......................................................176
TABLE 6. Target Item Ease Index for French-speaking World Cloze Test............180
TABLE 7. Target Item Ease Index for Muslim World Cloze Test........................181
TABLE 8. Raw Pretest Target Form Accuracy Scores (out of 24) .........................185
TABLE 9. Estimated Marginal Means and Standard Errors for Immediate Posttest Target Form Accuracy by Corrective Feedback Type.................................187
TABLE 10. Estimated Marginal Means and Standard Errors for Immediate Posttest Target Form Accuracy by Input Enhancement Level.................................187
TABLE 11. Estimated Marginal Means and Standard Errors for Immediate Posttest Target Form Accuracy by Corrective Feedback Type by Input Enhancement Level...............................................................188
TABLE 12. ANCOVA Results for Immediate Posttest Target Form Accuracy.........190
LIST OF TABLES—continued

TABLE 13. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Corrective Feedback Type ..........................192

TABLE 14. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Input Enhancement Level ............................194

TABLE 15. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Corrective Feedback Type by Input Enhancement Level ..................................................195

TABLE 16. ANCOVA Results for Delayed Posttest Target Form Accuracy ...............197
ABSTRACT

Despite popular theoretical arguments against the usefulness of explicit knowledge in the development of linguistic competence (cf. Krashen, 1982; Schwartz, 1993), empirical studies have suggested that comprehensible input alone may be insufficient in fostering target-like L2 production skills (Swain, 1985; Hammerly, 1987; Klein & Perdue, 1997). As such, many SLA researchers (e.g., Long, 1991; 1996; 2007; Spada, 1997; Lyster, 1998; 2001; 2004) now advocate teacher use of negative feedback during communicative activities to promote learner noticing of errors and internalization of correct forms; yet the efficacy of this practice remains questionable as it not only entails the provision of additional input but generally enhances otherwise non-salient linguistic features as well (Leeman, 2003). Moreover, these potentially confounding variables may be particularly problematic with regard to the acquisition of grammatical structures bearing little informational content such as the à/au/en/aux distinction before French toponyms.

In a controlled experiment with a pretest-posttest-delayed posttest design, all French 102 students at the University of Arizona in Fall 2006 were randomly assigned to an interactive, computer-based reading lesson featuring either typographically enhanced or unenhanced target forms and one of four types of corrective feedback: meaning-focused,
implicit form-focused, simple explicit form-focused, or explicit form-focused with rule explanation. Statistical analyses of the results on both the short-term and long-term linguistic outcome measures showed continuous improvement in target structure accuracy for learners in all treatment conditions. However, there was no significant advantage for any of the investigated feedback types, nor for target form enhancement, and no interaction between these two variables.

This study assessed learners’ pre-treatment performance on the target structures but not their explicit prior knowledge of them. Thus, it is possible that a fine-tuning of the latter through exposure to comprehensible input reduced the magnitude of the learning effects, especially given the proximity of the target prepositions to their objects. Nevertheless, these findings suggest that neither highlighting nor repeated correction and explanation of these particular grammatical items yields any additional benefit and that instructional efforts may better be focused on maximizing beginning learners’ opportunities to process authentic language through meaningful activities.
CHAPTER 1
INTRODUCTION

Debates over language teaching methodology are probably as old as methods themselves, and although these discussions today are largely framed in terms of ‘practices and principles’ (Richards & Rodgers, 2001), claims for the efficacy of one method over another are no less hotly contested. Central to most arguments are the utility or futility of explicit instruction and correction. Proponents of the latter view (e.g., Krashen, 1982; Pinker, 1991; Schwartz, 1993) contend that first and second language acquisition are analogous processes controlled by a natural human endowment, known as Universal Grammar (UG), which matches linguistic data from the environment to an innate store of possible structural rules to determine those that are operational in the given language and then generate novel utterances accordingly. Advocates of the former, on the other hand, generally either reject the idea of UG altogether (e.g., N. Ellis, 2002b) or distinguish first and second language acquisition as being fundamentally different (e.g., Bley-Vroman, 1990). While each side has attempted to provide evidence in its own favor, no clear-cut victor has so far emerged as the results of many published studies to date are subject to interpretation if not flawed in design or execution. In the meantime, however, language teachers are left to wonder whether their
classroom activities and behaviors are having the intended effects, for it is commonly believed that unnecessary correction can raise learner anxiety and, thus, inhibit cognitive processing (VanPatten, 1989), whereas ignoring learner errors may result in fossilization, that is, the permanent stabilization of nontarget-like forms (Higgs & Clifford, 1982).

The present investigation grew out of an attempt to solve a very practical problem, namely the pedagogical handling of a set of high-frequency French grammar structures—the prepositions à, au, en, and aux with geographical place names—in view of previously observed evidence of their imperfect acquisition even after explicit rule presentation. Through the results of a tightly controlled computerized laboratory experiment, this study aims to shed additional light on the contentious issue of grammatical correction and suggest implications for classroom language instruction.

1.1 Historical and Theoretical Background of the Problem

Although most people may not question the old, familiar adage that “we learn from our mistakes,” perhaps nowhere does this idea seem to generate more controversy than in the field of second language acquisition (SLA). Consequently, over the past half-century, the pedagogical pendulum has swung from the highly corrective form-focused method of Audiolingualism to the more permissive meaning-based approach of Communicative
Language Teaching and now back again to Focus on Form. As last year’s ‘best practice’ becomes this year’s worst, classroom teachers must feel a strong sense of frustration if not professional vertigo. To break the pattern of *va-et-vient*, however, will require a closer examination of the chain of ideas and events that have led to this vacillation.

While correction in foreign language instruction can trace its origins to the earliest days of Grammar Translation, it achieved a special prominence in the middle of the 20th century with the advent of the Audiolingual Method (ALM), the outgrowth of the Behaviorist learning theory that dominated American psychology at the time. Although it was actually Pressey (1926) who first emphasized the importance of immediate feedback in shaping learner responses in education, B.F. Skinner’s (1957) characterization of language as the product of habit formation was what gave rise to ALM. In Skinner’s view of learning in general, behavioral changes take place as a result of consequences, which, depending on their desirability, either increase or decrease a behavior’s frequency (cf. Skinner, 1953). Thus, positive feedback (such as praise) will promote recurrence, and negative feedback (a.k.a., punishment, such as criticism) will discourage it.

In fairness to Skinner, it should be pointed out that he questioned the efficacy of negative feedback and strongly advocated the division of complex learning tasks into very
small steps, the successful completion of each of which to be positively reinforced. He further suggested that “by making each successive step as small as possible, the frequency of reinforcement can be raised to a maximum, while the possibly aversive consequences of being wrong are reduced to a minimum” (Skinner, 1968, p. 21).

Although this type of instructional design is easier imagined than accomplished, especially when dealing with learners in groups rather than as individuals, from this statement alone, one might still wonder how corrective feedback might ever have become associated with Skinner. However, elsewhere in the same volume, Skinner concedes that contingencies of reinforcement can be improved by punishing small units of behavior, such as through mild reprimands. In fact, coincidentally with respect to the development of writing skills, he proposes that “a series of small punishments for bad grammar, illogical constructions, and solecisms, for example, may be useful” (Skinner, 1968, p.187). Thus, particularly when viewed in light of the long history and natural tendency toward aversive control in education in general (cf. Skinner, 1968), the eventual distortion of Skinner’s ideas is certainly understandable. When applied to language learning, the common interpretation of Skinner’s theory implied that target-like linguistic performance resulted from contingencies of interlocutor approval and disapproval (Brown & Hanlon, 1970).
Although his intent at the time was merely to explain first language acquisition, Skinner’s (1957) assertion that all types of learning were shaped by the same mechanisms obviously pertained to SLA as well. By analogy with behavior modification, L1 linguistic patterns came to be viewed as habits to be unlearned and replaced by those of the L2. Thus, the classroom practice of ALM consisted largely of mechanical drills, in which the learners listened to model dialogues featuring key language structures and then repeated them line by line while the teacher corrected their pronunciation and grammar mistakes explicitly and immediately until their delivery was perfect (Richards & Rodgers, 2001).

After a decade of preeminence in both foreign language classrooms in the US and English as a second or foreign language (ESL/EFL) classes worldwide, ALM gradually fell into disfavor as it failed to foster a transfer of skills to situations of real communication (Richards & Rodgers, 2001). During this period, a critical mass had also developed for the new mentalist paradigm of formal linguist Noam Chomsky, who vigorously challenged Skinner’s characterization of verbal behavior on numerous grounds (cf. Chomsky, 1959). Notable among his objections was Skinner’s failure to explain what Chomsky would later call (linguistic) *competence*, an internalized system of rules allowing the speakers of a language to produce and understand an unlimited number of grammatical utterances outside
their experience. In his alternate account of the language learning process, Chomsky (1965) postulated the existence of a Language Acquisition Device (LAD), a ‘black box’ in the mind that would generate an abstract grammar based on the linguistic data it received (i.e., the language in the surrounding environment). Although the LAD has since morphed into the principles and parameters of Universal Grammar (Cook, 1993), the theoretical existence of an innate human capacity specifically for language acquisition enjoys an unrivalled status in the debate on both first and second language development to this day. In fact, Eubank and Gregg (1995) have even gone so far as to claim that the acquisition theory elaborated by the UG framework is “the only one there is” (p. 51).

Granted that UG may have provided an attractive alternative to the Behaviorist theory of language learning, it maintained the prevailing structuralist view of the nature of language itself. On the other hand, American sociolinguist Dell Hymes (1972) introduced the idea of communicative competence to expand on Chomsky’s basic concept and include knowledge not only of a given utterance’s grammaticality but of its appropriateness in any given context as well. Meanwhile in Europe, linguist D.A. Wilkins devised his own definition of language as communication, replacing grammar and vocabulary with the notional categories (e.g., time, quantity, and frequency) and communicative functions (e.g., apologies, requests, and offers).
that would inform his *Notional Syllabuses* (Wilkins, 1976).

Thus, the next wave in L2 pedagogy—Communicative Language Teaching (CLT)—grew out of a new theory of language as much as one of learning. In fact, with regard to learning, CLT was based less on a formalized theory than on three main principles: Activities should 1) involve real communication 2) for the completion of genuine tasks 3) using language that is personally meaningful to the learners (Richards & Rodgers, 2001).

Admittedly, subjective descriptors such as *real, genuine,* and *personally meaningful* are difficult to qualify, but an example of an activity that fares poorly on all three of these criteria would be one in which the teacher dictates the language structures or vocabulary to be used in an exchange motivated only by his or her directive undertaken for the sole purpose of satisfying the requirements of a prescribed assignment, such as an information gap exercise in which the learners must obtain information about a list of fictional characters by asking personal questions (e.g., Where is X from?) and filling in a table (Lyddon, 2003).

With CLT’s emphasis on the primacy of communicative meaning over grammatical form, the efficacy of corrective feedback in the L2 classroom started to come into question, thus forging a new nexus with the Chomskyan generativists, for a principle tenet of UG is that language acquisition occurs solely on the basis of positive evidence, that is, information
as to what utterances in a language are grammatical. Corrective feedback, on the other hand, falls under the umbrella of negative evidence, or information as to what utterances a language considers ungrammatical.

Pinker (1991) argues that since L1 acquisition is universally successful in all normally developing children, in order for negative evidence to play any significant role in the process it would not only have to be available to children of all first languages, but it would have to be useful, used, and necessary. According to what Pinker calls the ‘blame assignment’ problem, even if all children were exposed to corrective information, they may not always be able to identify the error that solicited it. Furthermore, Pinker contends that in cases of overt correction, children often either resist or ignore it as illustrated by the following oft-cited exchange between a mother and son:

Child: Nobody don’t like me.
Mother: No, say “nobody likes me.”
Child: Nobody don’t like me.
    
    
    
    
    (eight repetitions of this dialogue)
Mother: No, now listen carefully; say “nobody likes me.”
Child: Oh! Nobody don’t likes me. (McNeill, 1966, p.69)

Finally, Pinker claims that negative evidence is clearly not necessary as adults are able to judge the ungrammaticality of sentences they have never even uttered much less had corrected.

Assuming that first and second language acquisition both rely on the same underlying processes, if first language acquisition requires positive evidence alone, corrective feedback would naturally be unnecessary in SLA as well. One of the best-known and most outspoken proponents of this view has been applied linguist Stephen Krashen, who posits a distinction between the implicit process of language *acquisition* and the explicit process of language *learning*. His basic position consists of three main tenets: 1) that acquisition is the primary mechanism in developing linguistic competence, 2) that consciously learned linguistic knowledge can never become implicit, and 3) that the sole utility of explicit knowledge is to permit monitoring of language once it has already been formulated (Krashen, 1982). As such, he claims that comprehensible input is the sine qua non of SLA and that error correction, as it aims at rectifying explicit rather than implicit knowledge, makes little or no contribution to true linguistic development. In other words, knowing the rules of a language and knowing the language itself are two different things; thus, it is understandable that students may show
mastery of various linguistic structures on discrete-point achievement tests but demonstrate little or no improvement on integrated tasks incurring spontaneous language use. The practical outgrowth of Krashen’s (1981) Input Hypothesis was the Natural Approach (Krashen & Terrell, 1983), which maintained that if language teachers focused their instruction on meaningful input, their students would acquire the target language at a natural rate, and the formal aspects of the linguistic code would essentially take care of themselves.

At the time of their appearance, Krashen’s ideas found widespread appeal and became influential in teacher education programs and L2 classrooms throughout North America. Notable among the voices of dissent, however, were Theodore Higgs and Ray Clifford, whose landmark essay (cf. Higgs and Clifford, 1982) both challenged the tendency of American foreign language instructors to simply ask whether students can communicate in their L2 as opposed to how well and emphasized the roles that linguistic accuracy and cultural authenticity should play in any true assessment of communicative competence. Furthermore, it argued that many L2 learners exhibit fossilized language behaviors as a result of inadequate instruction from the beginner level.

Although Higgs and Clifford have been criticized for their lack of empirical data (cf. VanPatten, 1989), empathizers such as Ke (1995) have echoed their sentiments. In his case
study on intensive Chinese programs, Ke laments how even many highly motivated students
are unable to move beyond a particular proficiency level. While it should be pointed out that
he bases his assessment on the ACTFL Oral Proficiency Interview rating scales, which take
explicit account of grammatical accuracy and give increasing weight to it at progressively
higher levels, Ke blames his students’ lack of attainment on their failure to master certain
basic structures from the start. Moreover, he observes how students from communicative
high school programs enter university at a considerable fluency advantage over those who
are taking Chinese for the first time. Unable to eradicate their fossilized errors, however, they
end up becoming demotivated and drop out after the first semester, once the other students
have surpassed them in terms of accuracy.

Another of Krashen’s challengers has been Long (1983), whose review of 16 SLA
studies comparing the relative efficacy of natural exposure as opposed to formal instruction
called into question the supposed futility of the latter. While Long agrees that meaningful
input is essential to L2 acquisition, he also argues that it is insufficient. Data from the French
Immersion programs in Canada appear to corroborate these assertions, for they reveal stark
differences between L2 French learners’ receptive and productive abilities even after
prolonged exposure to the language (Harley & Swain, 1984; Swain, 1985; Hammerly, 1987).
Despite his detractors, Krashen has some influential allies who also still subscribe to the Identity Hypothesis, that is, that L1 and L2 are basically acquired through the same process (Saville-Troike, 2005). For instance, Schwartz (1993) suggests that full-fledged L2 competence may not be possible for learners beyond a certain critical period of biological maturation, for they may no longer have complete access to UG. Thus, the best these learners can do for some L2 structures is to consciously learn and practice them to the point of automaticity, which is still not technically the same as acquisition, though the two phenomena may resemble each other on the surface.

Not all theoreticians working within the UG paradigm believe that L1 and L2 acquisition are driven by the same mechanisms, however. Bley-Vroman (1990), for example, proposes the rival Fundamental Difference Hypothesis, which holds that L1 and L2 are, in fact, acquired through completely different processes. Within this view, L1 is constrained by UG and develops via a biologically adapted language-specific learning process, but as this route becomes unavailable with age, L2 development must rely on L1 and general problem-solving ability as surrogates. Essentially, language learners by default will assume that L2 operates by the same rules as L1. In cases when two parameters are mutually exclusive (e.g., head-first vs. head-last), positive evidence alone may be enough to trigger the abandonment
of this original assumption, but in instances where one parameter is a subset of another (e.g., pro-drop vs. non-pro-drop), negative evidence is necessary to allow a reduction of the learner’s hypothesis space.

Negative evidence may be either preemptive (e.g., linguistic rules or patterns taught prior to communicative activity) or responsive (e.g., corrective feedback given in the context of communication), but the availability of any linguistic evidence, whether positive or negative is no guarantee of its internalization. In fact, Schmidt (1990, 1993) argues that the relevant features of the linguistic input must be consciously noticed in order to become incorporated into the learner’s interlanguage. An important question is how to draw attention to these features within the context of real communication.

Illustrating the problem is VanPatten’s (1990) study of 1st to 3rd year university L2 Spanish learners, in which he found that asking them to focus on linguistic form and communicative meaning at the same time negatively affected their listening comprehension. It would then seem that if form-focused instruction is to take place during communicative activities it will have to do so at a point where the cognitive burden on semantic processing is otherwise light. As one possible solution to this apparent conundrum, Long (1991) proposes what he calls *Focus on Form*, or FonF.
Long and Robinson (1998) distinguish FonF from traditional form-focused instruction in that, despite the name, its true main focus is still meaning. In other words, the attention to form takes place in a context of purposeful communication (e.g., tasks) as opposed to mere language practice (e.g., pattern drills). In this respect, FonF is similar to Krashen and Terrell’s (1983) Natural Approach. A critical difference, however, lies in the way its practitioners handle learner production errors, of which Ellis, Basturkmen, and Loewen (2002) identify two types: conversational and didactic.

Conversational errors are those that adversely affect interlocutor comprehension and, thus, require attention to form to repair a breakdown in communication. Consequently, in response to learner utterances such as “Yesterday go shopping,” where the actor is unclear, FonF and Natural Approach proponents alike might model a target-like alternative while confirming the learner’s intended message with a reformulation like “Yesterday you went shopping?”

Didactic errors, on the other hand, do not obscure meaning. For instance, in the case of an utterance such as “Yesterday I go shopping,” the communicative intention is arguably clear. Here FonF advocates might still respond with the same reformulation as in the previous example, whereas those in favor of the Natural Approach would ignore the error
and perhaps continue the exchange with “Oh, really? What did you buy?” In short, FonF
allows for occasionally shifting learners’ attention to matters of form so long as meaningful
communication remains the principal objective of the overall activity.

Long (2007) advocates the use of only responsive negative feedback so as not to
unwittingly direct learners’ attention solely to grammatical form throughout the activity and,
thus, undermine its intended communicative purpose. More specifically, he endorses an
implicit feedback type known as recasts, by which the teacher repeats the learner’s utterance,
but in corrected form and with a rising intonation at the end. A problem with this error
correction technique, however, is that it resembles the asking of a confirmation question: If
the learner misinterprets it as such, he or she may respond with a simple affirmative and, thus,
defeat the purpose of the recast (Lyster, 1998).

One means of raising learner consciousness of formal linguistic properties is to flag
them in the input, as through a change in vocal pitch, stress, or volume (Sharwood Smith,
1991). As such, enhancing the corrective portion of recasts might resolve their ambiguity.
However, since recasts themselves are grammatical, they constitute evidence of what is
possible in a language and only imply what is impossible by their adjacency to a learner’s
ungrammatical utterance. Furthermore, although Krashen has questioned the role that
salience might play in SLA (Dulay, Burt, and Krashen, 1982), N. Ellis (2005), who shares Krashen’s belief in the basically implicit nature of language acquisition, notes that once the human brain becomes optimally tuned to perceive L1 input it may require enhancement of certain stimuli for them to initially register, regardless of the type of evidence in which they occur. Thus, while enhanced recasts have been shown to have a positive effect on SLA (Doughty & Varela, 1998), one could attribute these results to any one of three causal factors: negative evidence, enhanced input, or input per se.

1.2 Rationale for the Present Study

As the literature review in Chapter 2 will show, despite many previous attempts to provide empirical evidence for either side of the debate over grammatical correction, none has so far achieved irrefutable success. First of all, while the procedure for demonstrating that comprehensible input plus corrective feedback is superior (or inferior, for that matter) to comprehensible input alone is relatively straightforward, it is difficult statistically to argue that there is no difference between the two, and in terms of acquisition, those who advocate the exclusive use of positive linguistic evidence make no other claims about negative evidence than its lack of added value.
While some studies have produced statistically significant results indicating an effect for correction (e.g., Carroll & Swain, 1993), the majority of these have only measured short-term outcomes (e.g., from one to five weeks). More serious still is the nature of the instruments that have been used. In many cases, they have consisted of discrete-point, paper-and-pencil tests, rather than integrative tasks. As VanPatten (1989) points out, although form-focused instruction may have a positive effect on testing, it would be difficult to make the same generalization about communicative ability. However, given that learners who ever reach target levels of L2 proficiency are as anomalous as those who fail to do so in their L1 (Bley-Vroman, 1990), it is imperative to continue investigating this issue, for comprehensible input alone (at least in the quantity that most classroom learners receive) clearly does not appear to be adequate. Long (2007) asserts that even if negative evidence is not absolutely necessary for language acquisition to occur, it may at least play a facilitative role, and in view of the very limited number of contact hours most language teachers have with students, further exploration of its use would be worth pursuing.

In addition to the need for longer-term studies with integrative measures, however, is the obligation to tease apart the various aspects of corrective feedback to discover not only if it is effective but why. As corrective feedback provides not only negative but also positive
evidence and potentially highlights structural features of the linguistic input as well, it is crucial to determine exactly which of these factors is responsible for any improvement in language acquisition, for non-corrective intervention, if equally effective, would be preferable as it eschews the risk of souring learner affect or lowering motivation.

Thus, the following study is an attempt to redress some of the shortcomings of the previous research on corrective feedback and thereby help clarify L2 classroom teachers’ options for handling learners’ didactic errors. Through random assignment of participants to computerized treatment conditions involving a communicative task and multiple versions of an integrative outcome measure in a pretest-posttest-delayed posttest experimental design realized over a 16-week period, this investigation managed not only to strictly regulate the quantity and salience of positive and negative evidence relative to the target structures but to determine the explicitness and elaborateness of the correction as well. While a laboratory-based study such as this one may lack the ecological validity of a classroom setting, this shortcoming is offset by superior control of extraneous variables as a result of random assignment to the experimental conditions of interest here. Finally, by choosing from a widely adopted introductory commercial coursebook a high-frequency but difficult-to-master grammar structure—the à/au/en/au distinction before French toponyms—it is hoped that
classroom French instructors will immediately recognize the relevance of this study to their daily teaching.

1.3 Definitions of Important Terms

A number of essential terms will recur frequently throughout this dissertation. While some of these have already been mentioned in brief and others will receive additional attention in later chapters, virtually all have taken on slightly different meanings in the works of various authors. As such, to clarify any misconceptions to this point and avoid any eventual further misinterpretation, the following paragraphs offer for convenient future reference a few precisions on how these key terms should be understood in this particular context.

First is the important distinction between positive and negative linguistic evidence. Simply put, positive evidence provides information that a linguistic form or structure is grammatical in a language, whereas negative evidence signals that one is ungrammatical (Pinker, 1991). As Sharwood Smith (1991) explains, the former “requires nothing on the part of the provider except to speak or write according to the native-speaker norms,” while the latter “always requires extra information” (p. 123). Gregg (2001), however, notes that as these terms were originally coined to describe factors in L1 acquisition, they do not neatly
account for two important phenomena prevalent in (if not exclusive to) SLA, namely metalinguistic explanation and corrective feedback. Thus, he suggests an alternate conceptualization of positive evidence as *language used* (i.e., in a non-metalingual context) and negative evidence as *language mentioned* (i.e., at least partly metalinguistically) (p. 170). Although some readers may argue the possibility of metatalk on a given structure simultaneously serving as positive evidence for its grammaticality, as the ostensible purpose of this type of discourse would be to provide learners with declarative knowledge for use in later monitoring (i.e., checking for errors in their own language performance), it also arguably constitutes explicit negative evidence and, thus, is considered as such in the pages of this manuscript.

Another important point for clarification is the difference between *meaning-focused* and *form-focused* instruction. Whereas the former requires learners to attend merely to the content of their communicative output, the latter attempts to draw attention to certain aspects of the linguistic code (R. Ellis, 2000). Representative of meaning-focused instruction (MFI) is Krashen and Terrell’s (1983) Natural Approach. Form-focused instruction (FFI), on the other hand, can further be subcategorized into three types: 1) focus on forms, 2) planned focus on form, and 3) incidental focus on form.
The primary aim of focus on forms (FonFs) instruction is to provide learners with intensive practice on pre-selected linguistic features. While FonFs activities may take either a ‘didactic’ (e.g., explicit, deductive rule presentation) or ‘discovery’ (e.g., implicit, inductive exemplar examination) approach, their emphasis on linguistic structure is readily apparent to both teacher and student (R. Ellis, 2000). Examples of this type of instruction include Grammar Translation and the Audiolingual Method.

Planned and incidental focus on form (FonF), on the other hand, both derive from a single term first proposed by Long (1991), whose original definition included only the latter type, that is, attention to linguistic features occasioned by communicative need (pp. 45-46). As R. Ellis (2000) points out, however, Long and others (e.g., Long & Robinson, 1998) have since extended use of the term to cover the identification of specific problematic forms and the development of subsequent activities to elicit them, presumably to facilitate research. In fact, it is this latter, planned variety that informs the form-focused treatment conditions in the present study.

Even within planned FonF, however, there are two delivery options. The first of these is preemptive FonF, which attempts to sensitize learners to certain correct linguistic forms before errors can occur (cf. Ellis, Basturkmen, & Loewen, 2002). Common not only to
planned FonF but to FonFs instruction as well, the important difference being the communicative nature of the activities that follow, preemptive teaching practices include both didactic and discovery rule awareness-raising techniques such as those mentioned earlier above. The product of this type of intervention is referred to in this dissertation as preemptive negative evidence.

The second planned option is to employ reactive FonF, which addresses learner errors only once they have been committed, in the context of communicative activity (cf. Ellis, Basturkmen, & Loewen, 2002). Thus, this type of FonF is synonymous with corrective feedback, which this dissertation will broadly call responsive negative evidence.

Corrective feedback, however, falls along a continuum ranging from explicit to implicit. Explicit corrective feedback overtly indicates that some aspect of a learner’s linguistic behavior is inconsistent with the norms of the target language, whereas implicit corrective feedback requires learner inference of such discrepancies (cf. Carroll & Swain, 1993). On the more explicit end of the spectrum are formal restatements of malformed utterances as well as metalinguistic explanations of appropriate alternatives. Examples of more implicit corrective moves then include clarification requests (i.e., prompts for repetition or paraphrase), confirmation checks (i.e., yes/no questions consisting of echoes or
paraphrases with rising intonation), or recasts (i.e., reformulations of all or part of individual learner utterances containing at least one error and occurring within the context of communicative activity; cf. Sheen, 2006).

The final terminological issue requiring clarification involves the concepts of salience and input enhancement. The first of these is notoriously difficult to define. Dulay, Burt, and Krashen (1982) suggest that salience consists of “particular characteristics that seem to make an item more visually or auditorily prominent than another” (p. 33). While acknowledging that the inherent qualities (e.g., duration, isolation) of certain stimuli may cause them to stand out more than others, the usage of the term in this dissertation will reflect an absolute rather than a relative criterion. In other words, any linguistic feature exhibiting properties sufficient to trigger conscious perception (i.e., awareness) is considered salient.

Input enhancement, then, is an attempt to generate salience externally for linguistic features that lack it intrinsically (Sharwood Smith, 1991). Strictly speaking, all form-focused instructional techniques constitute input enhancement. However, except where otherwise noted, in this dissertation the term will be employed in a restricted sense to exclude the explicit types such as corrective feedback moves and encompass only modifications to the physical forms of linguistic features prior to their initial appearance in a communicative
context. Furthermore, as it is not the entirety of the input but only certain key parts of it that
are highlighted, the alternate term target form enhancement will also be used interchangeably.
In any case, examples of such enhancement would include special stress, intonation, and the
use of gestures in oral modality and boldface, underlining, and color contrasts in written
mode. As the experimental treatment in this study consisted of a series of written texts, input
enhancement here refers specifically to typographical alterations of the latter type.

1.4 Outline of the Dissertation

The basic foundation for the present study now laid, below is an overview of the
remaining chapters of this manuscript. Chapter 2 provides a critical review of the relevant
literature informing the debates on the major topics of this introductory chapter, including the
benefits of formal instruction over natural exposure, the need for negative evidence, the
degree to which correction needs to be explicit, and the effectiveness of input enhancement
in promoting learner noticing of linguistic forms.

Chapter 3 gives a detailed account of the adopted research methodology. Following
an illustration of the experimental design, a statement of the research questions and
hypotheses, and descriptions of the participants and target linguistic structures, it elaborates
the outcome measures and exemplifies the instructional materials used in the treatment. It
concludes by outlining the data collection and handling procedures.

Chapter 4 then examines each of the research questions in light of the results of the linguistic outcome measures and explains the statistical decisions with respect to the corresponding hypotheses.

The body of the dissertation concludes with Chapter 5, an interpretation and discussion of the empirical results and their implications for L2 pedagogy as well as an acknowledgement of the limitations of the present study and suggestions for future research. Finally, an abundance of supplementary material, including copies of all the instructional materials and outcome measures is appended at the end.
CHAPTER 2

LITERATURE REVIEW

Whereas the preceding chapter presented a general overview of the problem under investigation, this chapter provides a more detailed examination of the theoretical issues and relevant empirical findings that shaped the eventual research design. It begins with a broad look at the benefits of L2 instruction as opposed to mere language exposure and then sequentially focuses on several of the most pertinent potential contributing factors, namely different types of linguistic data (i.e., positive and negative), different levels of interventional explicitness, and different degrees of target form salience. Finally, it concludes with a brief summary of the implications of the past research, especially with respect to the unanswered questions that the present study seeks to address.

2.1 The Benefits of Instruction

In light of the many failed past attempts to demonstrate the superiority of one L2 teaching method over another, Kennedy (1973) has noted the difficulty of making methodological comparisons. One problem has been that L2 instructors generally do not adhere to any single set of prescribed guidelines; thus, two teachers claiming to endorse very
different methodologies may in fact exhibit the same general practices in the classroom (Swaffar, Arens, & Morgan, 1982). However, given that most L2 learners never reach target-like proficiency levels (Bley-Vroman, 1990), the principles underlying these practices perhaps need to be examined more closely. As Bloomfield (1992) laments, there is not enough time in the average L2 classroom for all learning to occur by ‘osmosis’. In short, some sort of structured or systematic instructional approach seems necessary.

Dulay, Burt, and Krashen (1982), however, hold that language acquisition follows a natural, fixed developmental sequence that no manipulation of linguistic data can undermine. Thus, for Krashen (1981, 1982), the sole utility of language instruction is to provide learners with comprehensible input and opportunities for acquisition. In challenge to this position, Long (1983) reviewed 16 L2 studies that had been done by that time (mostly in English as a Second Language, or ESL) and concluded that instruction comprised more than mere language exposure. Yet a closer look at the studies he examined raises doubts about his interpretations.

In his review, Long (1983) compared the outcomes of varying amounts of classroom instruction and natural language exposure, starting with studies of learners receiving either both or only the latter. These covered all ages and proficiency levels, including beginning
through advanced ESL elementary and secondary school learners in mainstream classrooms
and pull-out programs (Hale & Budar, 1970; Fathman, 1975) and intermediate and advanced
ESL university students with or without formal language instruction (Upshur, 1968; Mason,
1971). Furthermore, these researchers employed a variety of discrete-point and integrative
instruments measuring language growth over periods of seven weeks to three years. Only
Hale and Budar, however, found any difference, in favor of exposure alone. While Long
legitimately points out that this conclusion was based on raw scores (as opposed to inferential
statistics), he can only rightly claim that those learners receiving classroom instruction did
not do significantly worse, not that they did better.

Corroborating the finding of no beneficial effect for instruction are two additional
studies, one of beginning through advanced elementary and secondary ESL learners
receiving 3, 5, or 10 hours of formal language instruction per week (Fathman, 1976) and
intermediate and advanced ESL university students receiving one or two hours of such
supplemental language support (Upshur, 1968). In short, these studies showed no advantage
for extra classroom time as opposed to out-of-class language exposure.

The picture becomes a bit more complicated, however, in light of two studies of
beginning through advanced adult ESL learners with several years’ residence in the US and
similar amounts of natural exposure but differing amounts of instruction or the reverse (Krashen & Seliger, 1976; Krashen, Seliger, & Hartnett, 1974). Predictably, those with more instruction showed higher levels of proficiency. Less easily explicable, though, is the discovery of no significant difference for greater amounts of natural exposure. St. Martin (1980), on the other hand, found that short-term intermediate and advanced adult ESL exchange students who participated in a homestay consistently performed better than their counterparts who lived in a dormitory.

Further clouding matters, however, are the last four studies in Long’s review, all of which examine the relationships between language proficiency level and respective amounts of instruction and exposure. A study of intermediate and advanced adult ESL learners (Krashen, Jones, Zelinski, & Upsrich, 1978) and one of beginning child learners of L2 Spanish (Brière, 1978) showed significant correlations for both criterion variables, but a stronger relationship for instruction. In a study of beginning through advanced university learners of various L2s in the US and abroad, however, Carroll (1967) found exposure to be the more strongly correlated. Finally, in the case of beginning through advanced adult L2 English learners in Japan, Chihara and Oller (1978) found a positive relationship only for instruction. Here, however, it is necessary to point out that for these learners, classroom
instruction was virtually the only English exposure they had. Judd (1981) distinguishes between acquisition-rich second language environments, where learners spend the majority of their time outside of class using the language, and acquisition-poor foreign language environments, where the language serves little communicative purpose outside the classroom. While researchers in the field of SLA do not generally maintain this distinction, it seems pertinent to this discussion, for in light of the mixed findings for groups of all other characteristics in the preceding studies, if indeed instruction offers any advantage, it is most likely among beginning adult learners in the latter type of setting (e.g., those in Chihara & Oller, 1978), who would otherwise have very little L2 exposure.

In addition to the provision of more linguistic input in general, another possible benefit of classroom instruction is that it may expose learners to structures that are more frequent in formal discourse. In a study of Italian ESL learners’ acquisition of relativization, Pavesi (1986) found a greater incidence of genitive clauses (e.g., “the woman whose children you met”) among learners with formal instruction than those without.

One commonality among the studies in Long’s analysis that may account for the relatively unimpressive case for instruction is their broad focus on acquisition rate and ultimate attainment. Pica (1983), on the other hand, in a study of L1 Spanish-speaking
learners of L2 English in classroom, naturalistic, and mixed settings, looked at a variety of specific morphological structures, namely those used in the acquisition order studies of Brown (1973) and Dulay, Burt, and Krashen (1982): progressive –ing, plural –s, singular copula, progressive auxiliary, articles, past irregular, past regular, third person singular, and noun possession –’s. All three groups acquired these morphemes in roughly the same order. Moreover, all exhibited both oversuppliance in nonobligatory contexts (e.g., “He studied English right now”) and omission in obligatory contexts (e.g., “He study English last year”). However, instructed learners showed significantly more overuse, whereas non-instructed learners tended toward underutilization. One interpretation would be that the two groups suffer from transfer-of-training (cf. Selinker, 1972) and failure to notice, respectively.

While overuse is arguably no better than underuse, Pica’s findings seem to support the Pedagogical Grammar Hypothesis (Rutherford & Sharwood Smith, 1985) that instructional strategies focusing specifically on L2 structural regularities can significantly accelerate their rate of acquisition beyond that which might be expected under natural circumstances where attention to form may be minimal and sporadic. Illustrating the natural rate of acquisition in the extreme is Klein and Perdue’s (1997) longitudinal study of uninstructed L2 learners of English, German, Dutch, French, and Swedish, which revealed
that about one-third of them indefinitely stabilized at a predictable nontarget-like linguistic system that the researchers dub the ‘Basic Variety’, which comprises mainly ‘open-class’ (i.e., lexically meaningful) words but no complementizers or inflectional morphology (e.g., markers of agreement, tense, aspect, or case).

The inadequacy of natural exposure, however, does not logically imply the superiority of instructional intervention. Harley (1989), for example, found that pre-teaching the use and formation of the *imparfait* and *passé composé* to 6th grade L2 French immersion students led to improved performance on only two of her three outcome measures in the week immediately following the treatment. Moreover, these differences disappeared by the end of three months. On a positive note, close inspection reveals that this last result seems to be due to gains made by the control group, as opposed to backsliding on the part of the treatment group, although the researcher did not provide data for any statistical significance tests for this possibility.

Day and Shapson (1991) employed a similar type of treatment, providing one of two groups of 7th grade L2 French immersion students with a special curricular unit including formal explanation of the use and formation of the conditional as well as form-focused activities to elicit its production. The researchers found that the learners in this condition
outperformed those in their control group on both an immediate cloze posttest and a written composition but not in an oral interview. Furthermore, these results were sustained in a second testing 16 to 18 weeks later. Besides the lack of differences in oral production, however, another problem with this study is that it fails to isolate the potential causative variables for the desired results. In short, the experimental treatment sensitized the learners to the target forms, provided them with additional exposure to them, and elicited their production, any or all of which might have contributed to superior performance.

A similar problem can be seen with Doughty’s (1991) study of intermediate ESL learners’ acquisition of relativization. In a computerized experiment, learners in three treatment conditions read a story presented as a sequence of single sentences, each containing a relative clause. While those learners who were shown the rule-oriented and meaning-oriented presentation both outperformed the control group on the immediate post-treatment measures, they also both received double exposure to the input as they saw the relative pronouns not only in their original contexts but in a subsequent animated mini-lesson as well. More troublesome, though, is the fact that the meaning-oriented group scored nearly twice as high as the other two groups on their comprehension of the story. In other words, learners in the rule-oriented group showed equivalent growth on the target structures despite the relative
incomprehensibility of the input, which calls into question the value of the outcome measures. Incidentally, those in the control group likewise evinced significant improvement from pre- to post-treatment, only not of the same magnitude. Unfortunately, their results may have been subject to ceiling effects because of their superior performance on the pretest.

In 2000, Norris and Ortega performed a meta-analysis of the research on the effectiveness on L2 instruction, which included 49 experimental and quasi-experimental studies published between 1980 and 1998. Considering only designs with treatments focusing on specific grammatical forms and dependent variables seeking to measure the effects on them (as in the three studies discussed immediately above), the researchers found a very large average effect size of almost one standard deviation (Cohen’s $d = .96$). In short, despite the lackluster look of some of the results described to this point, it would be hard to argue that instruction is completely without benefit. This construct, however, can consist of many components, including the presentation of rules, exposure to relevant input, opportunities for practice, and provision of corrective feedback, the first and last of which are the most disputable and which, thus, underpin the discussion in the next section of this review.
2.2 The Need for Negative Evidence

Although most classroom teachers have probably at some point given learners corrective feedback, this practice has long been a topic of controversy among SLA researchers. While opponents such as Truscott (1999) argue that it causes “embarrassment, anger, inhibition, and feelings of inferiority” (p.441), advocates might counter that it is like a bitter medicine: If it works, it should be prescribed and taken. The problem is a general failure to critically examine the implicit assumption in the lay use of the conditional. Feedback clearly plays an important role in shaping human behavior, but its relevance to language learning has been in question ever since Chomsky’s (1959) displacement of the Behaviorist construct of language with that of Universal Grammar (UG).

2.2.1 The Nativist Perspective

Within the UG paradigm, theoreticians speak of two types of linguistic data, often referred to as positive and negative evidence. Positive evidence consists of instances and exemplars of a language used in the context of communication and, thus, includes any non-metalinguistic statement a mentally competent adult L1 user makes. The importance of this type of linguistic data is that during acquisition it signals to the brain the syntactic constraints under which the target language operates. For example, a sentence such as “She plays the
“piano” permits the language learner to extrapolate that among the finite number of possible canonical word orders, SVO is the appropriate one for English. Once this constraint has been established, the learner can independently use it to generate grammatical utterances, including ones never previously encountered but which adhere to the same basic syntactic principle.

*Negative evidence*, on the other hand, comprises information about the impermissibility of some utterances, namely, ones that violate the underlying principles of the target language. In the case of L1 acquisition, the terms *negative evidence* and *corrective feedback* (or *negative feedback*) can be used interchangeably without incident. However, it will be important later to maintain a distinction in SLA, for the latter is technically a subset of the former. *Negative feedback* is reactive, whereas *negative evidence* may also be preemptive, as in the case of pre-taught rules (Long, 2007). In any case, for both first and second language acquisition, advocates of a need for negative evidence ask how the learner without it should know not to overgeneralize certain rules. For instance, while adverb movement in English may seem to be free, compare the sentences in (1):

1) a) Often she plays the piano.  
   b) She often plays the piano.  
   c) *She plays often the piano.*  
   d) She plays the piano often.
Clearly, there is in fact a subtle constraint that most non-linguists are probably not even aware of.

Nevertheless, UG theoreticians (e.g., Grimshaw & Pinker, 1989; Pinker, 1991) generally discredit negative evidence as playing any role in language acquisition. To explain why such rules do not become overgeneralized, UG invokes what is known as the Subset Principle, which states that the learner will always choose the smaller of two possible grammars (Berwyck, 1985). Thus, because the child L1 English learner never hears “She plays often the piano” in the input, he or she assumes that it is not allowable.

Pinker contends that if any form of negative evidence were effective it would most likely be parental feedback on children’s grammatical correctness such as expressions of approval or disapproval, but L1 researchers (e.g., Brown & Hanlon, 1970; Braine, 1971) have investigated this possibility and found no contingency. In other words, parents reinforced grammatically incorrect utterances just as often as correct ones, positive evaluation depending on propositional content rather than structural accuracy. However, approval and disapproval are not the only sources of negative evidence. Consider the following illustration of ‘approval’ in this exchange taken from Brown and Hanlon (1970):

Child:   Draw a boot paper.
Mother:  That’s right. Draw a boot on paper. (p.49)
Whether the child noticed the suppliance of the additional preposition in his mother’s response, of course, is unknown, but there is certainly implicit evidence of a problem with his utterance. In fact, parents are more likely to repeat their children’s language exactly when it is correct, whereas they tend to reformulate it or request clarification in cases of single errors (Bohannon & Stanowitz, 1988).

Even if negative evidence is sometimes available, however, Pinker (1991) still argues the improbability of corrective feedback being provided to all children of all languages on all problematic forms at the precise moment of need. In other words, it has non-universal availability. By extension, he then posits that negative evidence cannot possibly be necessary. To illustrate, he opines as follows:

“Virtually every adult speaker of standard American English would judge the sentence ‘I ladled the floor with paint’ to be ungrammatical. Is that because at some point everyone has uttered these verbs in these contexts and benefited from negative feedback? If someone’s personal history had not included such events, would he or she find such sentences acceptable?” (p.14)

At the risk of sounding impudent, as I do not have any problem with the grammaticality of “I ladled the floor with paint,” I hope there is no cause for alarm. In all seriousness, however, it seems obvious that humans are able to generalize patterns and do not need specific instances of every verb in every context.
Whether negative evidence is available or necessary, however, Pinker further argues that it may not even be useful. In what he calls the ‘blame-assignment’ problem, he asserts that a single utterance contains far too many rules (e.g., phonological, lexical, morpho-syntactic, pragmatic) for the learner to effectively isolate the one in question. Finally, even if negative evidence met all three other criteria, Pinker points out that children still often ignore it, as the following dialogue between Braine (1971) and his two-and-a-half-year-old daughter illustrates:

Daughter: Want other one spoon, Daddy.
Father: You mean, you want THE OTHER SPOON.
Daughter: Yes, I want other one spoon, please, Daddy.
Father: Can you say ‘the other spoon’?
Daughter: Other … one … spoon.
Father: Say … ‘other’.
Daughter: Other.
Father: Spoon.
Daughter: Spoon.
Father: Other … spoon.
Daughter: Other … spoon. Now give me other one spoon? (pp.160-161)

Other research, however, suggests that the degree to which children ignore corrective feedback may be exaggerated. Farrar (1990), for example, analyzed the natural interactions of L1 mother-child dyads and found a significant positive relationship between the number of reformulations of plurals and progressives during his first observation and the level of their
acquisition at the time of his second. However, other grammatical morphemes correlated with different discourse types. In another study, Farrar (1992) further discovered that L1 children were two to three times more likely to imitate grammatical morphemes in recasts than in models. Farrar suggests that the effects of negative evidence may be selective, depending on a number of factors, including the learner’s attentional focus, memory limitations, and current linguistic system.

Subsequent research has corroborated the findings of Farrar. Saxton (1997) conducted an experiment with young children aged 4 years 9 months to 5 years 6 months to test their acquisition of nonce verbs with irregular past forms under modeling and recast conditions. His results showed that those who had received recasts were more likely to use the correct forms. Saxton has now formulated what he calls the Direct Contrast Hypothesis, whereby juxtaposition of a grammatical utterance with an ungrammatical one allows comparison of the two in a context of joint focus and, thus, permits retreat from overgeneralization of rules. However, he also notes that uptake is not perfect. In other words, children do not always overtly react to the feedback, although it is possible that they may still be internalizing it.

Whereas the previous study included only immediate post-treatment measures, Saxton, Kulcsar, Marshall, and Rupra (1998) conducted an experiment testing for the same
effects on nursery school children aged 3 years 8 months to 4 years 6 months over a five-week period and also found significantly better use of the correct irregular past tense forms. Furthermore, these children showed greater grammatical sensitivity on a cleverly adapted grammaticality judgment test.

2.2.2 The L2 debate

Despite multiple findings contradictory to the assumptions of UG, old beliefs die hard, and the parallelism between L1 and L2 traces its origins much farther back. Dulay, Burt, and Krashen (1982), for example, have long rejected any significant role for negative evidence in SLA, largely on the grounds that Pinker would later articulate. The evidence they cite in support of their position, however, relates mainly to children (who arguably learn differently from adults) or to writing (which involves more than the immediate processing in which the present study is interested). The one study they mention that does not fall into either of these two categories is Holley and King’s (1971) observational account of graduate student teachers of German performing sample lessons with university learners of various levels as part of a methods course. The researchers’ argument, however, was not that corrective feedback did not work but that it inhibited learners’ expression. They recommended increasing teacher wait time to 5-10 seconds before ‘cueing’ responses as a last resort.
Higgs and Clifford (1982) were among the first to publicly object to what they saw as the excesses of non-corrective instruction. In their classic paper, they captured the imagination of a sizable segment of the teaching population by raising the specters of ‘fossilization’ (cf. Selinker, 1972) and ‘terminal profiles’. In brief, left to their own devices, language learners were doomed to permanent stagnation at a markedly nontarget-like level of proficiency. Although Higgs and Clifford’s account was largely anecdotal and lacked empirical evidence to substantiate their claims, vindication came soon afterward with the publication of data from the Canadian French immersion programs, which revealed stark differences between learner receptive and productive abilities even after prolonged language exposure and, thus, suggested that comprehensible input alone may be insufficient to bring about target-like levels of output (Harley and Swain, 1984; Swain, 1985). Hammerly (1987) calculated that 12th grade L2 French immersion learners had roughly 7000 hours of language exposure by the end of a special multi-year program known as SLACC (i.e., Second Language Acquisition through Classroom Communication). Nevertheless, transcripts of a set of recorded interviews revealed errors in 40.3 to 72.4% of all these learners’ utterances. As such, he suggests alternatives to early immersion, including preliminary language learning strategies classes.
VanPatten (1989), however, is among those who have sought to avoid a pedagogical return en masse to the structural ways of the not-so-distant past. Not only does he mock Higgs and Clifford for introducing a state of ‘fossilophobia’, but he tries to restore order by rightly pointing out Harley and Swain’s (1984) explicit statement of having found “no evidence of fossilization in any particular L2 domain at any particular level” (pp. 301-302). Furthermore, he adds another facet to the argument against the usefulness of negative feedback by noting the multiple demands on learner attention (other than in mechanical drills) such as word choice and sequencing, pronunciation, comprehension of the reply, and 3rd party observation. He concludes that “in the best possible scenario, the average early stage language learner has difficulty attending to errors and error correction should we wish it” (p. 254).

Raising a number of other concerns with the provision of corrective feedback is Truscott (1999), who presents the issue from the perspective of the teacher as well as the learner. Beginning with the former, he contends that teachers often misunderstand the nature of learner errors, fail to provide feedback clearly and consistently, or lose their focus on communication. He then argues that learners in turn must not only notice the corrective signal but value it enough to consider its meaning, understand it, accept it, and try to
incorporate it into their interlanguage. Finally, he draws implications with respect to the learning-acquisition distinction: If the distinction is valid, learned knowledge can never become acquired, in which case corrective feedback is useless. On the other hand, if the distinction is invalid, controlled conscious processing may lead to automaticity, which would require motivated practice. Like Higgs and Clifford, Truscott has come under attack for not supporting his claims with empirical evidence (cf. Lyster, Lightbown, & Spada, 1999). Nevertheless, his arguments provide a nice outline of many of the research agendas that will follow.

Implicitly or explicitly, underlying the advocacy of positive evidence alone in L2 pedagogy is the assumption that first and second languages are acquired through the same processes, which is questionable in its own right. Although some learners do eventually attain native-like levels of grammatical knowledge, most do not (Bley-Vroman, 1990). Furthermore, the differences between L1 and L2 users’ grammatical competence cannot always be attributed to structures that UG would predict (Birdsong, 1992). As such, some researchers have suggested that the principles of UG may mature and become inaccessible after a critical period in L2 (cf. Schachter, 1996) just as they do in L1 (Lenneberg, 1967).

Bley-Vroman (1990) suggests that older learners may have to resort to L1 knowledge
and general problem-solving abilities in the absence of UG and their natural language
learning endowment. Consequently, they will erroneously hypothesize that the rules of their
L1 extend to their L2 when the former has a less restrictive grammar. For example, L1
French speakers will initially assume the grammaticality of English sentences like “She plays
often the piano” because French allows the adverb in the post-verbal position (i.e., Elle joue
souvent du piano), and negative evidence would be required to disconfirm this hypothesis
(Bley-Vroman, 1986). L. White (1987), who likewise questions whether UG is fully
accessible in late acquisition, seconds this opinion in her critique of Krashen’s Input
Hypothesis.

Schwartz (1993), however, argues that if SLA followed general hypothesis-testing
procedures such as those that Bley-Vroman describes, explicit data and negative evidence
should be more fruitful than they are. Yet this position assumes that learners are always
actively testing one or more hypotheses and are, thus, attuned to the feedback, whereas
researchers such as Schmidt (1990) emphasize the role of attention and noticing above all
else before any acquisition can take place. Moreover, Terrell (1991), who distanced himself
somewhat from Krashen after their collaboration on the Natural Approach (Krashen &
Terrell, 1983), later suggested that conscious monitoring of language use could interact with
acquisition, leading learners to acquire their own output, especially in foreign language contexts, where they are exposed to little input (cf. Judd, 1981). He then argued that these learners needed strict error correction to avoid wholesale acquisition of incorrect forms and structures. White, Spada, Lightbown, and Ranta (1991) echo this opinion, adding that learners in these contexts may also receive incorrect ‘positive evidence’ (i.e., nontarget-like input with no means of distinguishing it as such) from their classmates.

Incidentally, the phenomenon that Terrell describes can be observed even in input-rich environments, such as the one detailed in Schmidt and Frota (1986), where the first author recounts his experience learning Portuguese in Brazil. For instance, when his friend pointed out to him that he had mistakenly used the word marida for esposa, Schmidt felt sure that he had heard marida before. Close inspection of his diary entries later revealed his intuitions to be correct: He had said the word himself on at least three different occasions.

In light of the many apparent differences between L1 and L2 acquisition, Schachter (1991) urges that negative evidence not be tarred with the same brush as behaviorism. Whether or not negative evidence is actually necessary in SLA, the possibility remains that it may at least facilitate the process (Long, 1996, 2007). As such, many SLA researchers have formally investigated its usefulness.
2.2.3 Empirical Findings in SLA

At this point, it will be useful to revisit a point made earlier, namely the essential distinction between preemptive and responsive negative evidence. The preemptive variety attempts to predict learner errors and prevent their initial occurrence. Examples include didactic rule explanations as well as inductive ‘discovery’ techniques, the main criterion being that the instructional intervention precedes communicative use of the target forms.

Responsive negative evidence, on the other hand, provides information that an observed language-related behavior was anomalous, unacceptable, or otherwise unsuccessful. As such, this type of negative evidence is synonymous with corrective feedback, the definition of which adopted here is intentionally broad in order to cover responses not only to spoken and written utterances but to non-verbal signs of miscomprehension as well.

Corrective feedback is arguably more natural as it (at least sometimes) occurs for children acquiring their L1, whereas L1 rule teaching is reserved for prescriptive elements of formal discourse, which even monolinguals master to widely varying degrees. However, corrective feedback also carries with it a connotation of criticism for wrong-doing, which may negatively affect the recipient’s feelings. The following discussion will attempt to assess the relative merits not only of negative evidence but of its preemptive and responsive sub-
types as well.

### 2.2.3.1 General Effects of Form-focused Instruction

In a post hoc analysis of observational data on intact classes of 5th and 6th grade beginning-level ESL students in Québec, Lightbown and Spada (1990) compared the percentages of time spent per teacher on form-focused instruction and suggested a positive relationship with accuracy on certain morphological features (e.g., plural \(-s\), progressive \(-ing\)). However, the pattern is sometimes inconsistent and difficult to interpret because of possible effects of numerous confounding variables as well as an interaction with age. Furthermore, the specific grammatical features selected for analysis were not necessarily targeted by the instruction, which included both preemptive and responsive interventions. White, Spada, Lightbown, and Ranta (1991), on the other hand, looked at a similar population that had received both explicit instruction and corrective feedback on question formation and found after five weeks that these learners showed greater development on the linguistic target than did those who received no negative evidence.

Herron and Tomasello (1988; Tomasello & Herron, 1988, 1989) have demonstrated, however, that the timing of negative evidence can make a significant difference in its effectiveness. In all three studies, the researchers worked with beginning university L2
French learners and compared traditional rule teaching and modeling with a technique known as the Garden Path, wherein learners are first presented with a rule and then prompted to make overgeneralizations to anomalous cases so that they may be corrected. In each case, the researchers found that forcing learners to formulate a hypothesis and then disconfirming it through corrective feedback led to superior performance on immediate post-treatment written measures. Furthermore, the effectiveness of this strategy was evident on both new and review structures. Herron (1991) later showed that the Garden Path worked for oral as well as written measures.

A study by VanPatten and Cadierno (1993) involving 2\textsuperscript{nd} year university L2 Spanish learners, however, showed that intervention does not necessarily have to take place after production to be effective. To compare the effects of traditional rule-based teaching with grammatical correction and an approach known as Processing Instruction (PI), the researchers introduced the learners to OVS sentences with clitics (e.g., ‘La sigue el señor’, ‘Her-OBJ follows the man-SUBJ’). While learners in both instructional conditions were given explanations of the rules as well as opportunities for practice, the traditional group received corrective feedback in production, whereas the PI group received it during comprehension activities. Both groups showed significant gains for production, despite the fact that the PI
group had never practiced it. However, only the PI group improved in comprehension. It should be pointed out, though, that this approach can only work for structures that could cause communicative breakdown (i.e., conversational errors, as opposed to didactic).

L. White (1991), on the other hand, studied 5th and 6th grade francophone learners’ adverb placement in L2 English. Intent on testing the accessibility of UG in late language acquisition, she sought to determine whether negative evidence could trigger resetting of the parameter that allows verb-raising over the adverb in French (the literal equivalent of “She plays often the piano”) but forbids it in English. She found that the combination of prior explicit rule teaching plus corrective feedback led to significantly better results on her outcome measures than did comprehensible input alone. Although this effect lasted up to five weeks, it disappeared within one year.

More damaging to her conclusions than a lack of sustained improvement on her delayed posttests, however, is White’s failure to control the input. The treatment group received instruction on adverbs, while the control group studied question formation. White had assumed that the level of exposure to the target word-order patterns would be the same for both groups, but later discovered on examination of her audio recordings of their lessons that adverb frequency in natural conversation was very low. In other words, the control group
received a much lesser amount of positive evidence.

As none of White’s three measures of the independent variable (i.e., a grammaticality judgment task, a multiple-choice task, and a scrambled sentence task), required real-time processing of linguistic form during communicative activity, it is little wonder the treatment group outperformed the control. Form-focused instruction has been shown to have a positive effect on discrete-point grammar testing, but it would be difficult to extend that generalization to communicative ability (VanPatten, 1989). When she discovered that her learners had not maintained their immediate post-treatment gains in the long term, however, White tellingly suggested that “those exposed to negative input might have failed to retain knowledge that [the nontarget-like structure slated for eradication] is ungrammatical because of lack of suitable follow up or subsequent emphasis on this issue” (p. 159). Apparently, even she herself still harbors doubts as to the existence of any interface between learning and acquisition.

Adding insult to White’s injury, however, is a critique by Schwartz and Gubala-Ryzak (1992), who challenged her grasp of the UG parameter in question. Although their discussion is too complicated to enter into detail, in essence they argue that verb-raising is a two-step process that only becomes transparent in limited circumstances. Furthermore, verb-
raising is not only allowable but either required or prohibited, depending on the language, which is why SVAO sentences are acceptable in French and SAV(O) ones are not, whereas the reverse is true for English. Finally, positive evidence alone can indeed intimate which of the two possible parameter settings is operational in English, although such evidence is rare and not particularly robust.

In an attempt to salvage White’s earlier work, Trahey and White (1993) studied a single group of demographically similar learners acquiring the same parameter setting, only this time providing them with a two-week input flood of specially prepared materials modeling natural usage of English adverbs. Their results seemed to provide one missing puzzle piece as these learners increased their acceptance of the correct English SAV(O) order but failed to reject the incorrect SVAO. In short, they resembled those in the positive plus negative evidence condition for SAV(O) but those in the positive evidence only group for SVAO in White (1991). Moreover, in violation of the Uniqueness Principle (cf. Pinker, 1984), they indicated the possibility of simultaneously maintaining two supposedly mutually exclusive parameters.
2.2.3.2 Unique Effects of Corrective Feedback

While the studies mentioned to this point have all included both preemptive and responsive negative evidence (i.e., pre-production rule instruction and post-production corrective feedback), those discussed in this section seek either to separate the effects of the two or to examine those of the latter alone. VanPatten and Oikkenon (1996), for instance, looked to tease apart the effects of the rule explanations and structured activities that are part and parcel to Processing Instruction. Working with 4th semester high school L2 Spanish learners and the same OVS structures from VanPatten and Cadierno (1993), they compared groups receiving either PI (i.e., rule explanation plus structured input) or only one of its two component factors. The results showed that PI and structured input alone led to significantly superior performance as compared to explanation alone on an interpretation task, but only learners in the PI group similarly benefited in production. The researchers then concluded that the structured input, not the rule explanations, was the critical factor in determining the effectiveness of PI. As these structured input activities include corrective feedback, they also imply the relative value of responsive as opposed to preemptive negative evidence.

Whereas VanPatten and Oikkenon concerned themselves with intervention at the point of input processing, Lyster (2004) examined the effects of corrective feedback on
output. Using a quasi-experimental design, he measured 5th grade immersion L2 French learners’ developing accuracy of grammatical gender. Three of the four groups received preemptive negative evidence in the form of explicit instruction on gender cues in word endings. Furthermore, two of these three groups received corrective feedback via either prompts or recasts. The fourth group received no negative evidence at all. While oral and written posttests administered eight weeks after the treatment showed a general advantage for form-focused instruction, only the group receiving prompts demonstrated this effect consistently. Moreover, in some cases, the other negative evidence groups did not differ significantly from the control.

Doughty and Varela (1998) dropped preemptive negative evidence from the equation and looked only at the presence or absence of corrective feedback. In a quasi-experimental study, the researchers compared middle school-aged ESL students’ acquisition of past time reference in reported speech (e.g., “I thought that would happen” vs. “I thought that will happen”) in content-based science classes focusing only on meaning and those incorporating recasts during student presentations. Although the learners may have been exposed to explicit instruction on the target forms at some point previously, such intervention was not part of the research design. The researchers found improved oral and written use of the target structures
that was not only significant but sustained for two months.

One shortcoming of the preceding study, however, is that it does not account for possibly considerable differences in amounts of input of the target forms. Long, Inagaki, and Ortega (1998), on the other hand, controlled this potential confounding variable by comparing university learners in two separate experiments. Both designs comprised three treatment conditions: models, recasts, and a control. In the model condition, the learners heard the target structure and then had to repeat it. Learners in the recast condition had the target structures elicited and then corrected. Those in the control group only completed the outcome measures. One experiment looked at locatives, adjective order, and the conjunctor –kute in 2nd semester L2 Japanese, but found no differences for any treatment condition on the first two. Furthermore, the model and recast groups both performed significantly better than the control group on the –kute form, with no difference between them, implying that negative evidence may be neither necessary nor beneficial for its acquisition. The second experiment looked at adjective order and object topicalization in 3rd semester L2 Spanish. Here the results presented a slightly different picture: For adjective order, not only did recasts and models both lead to superior performance, but the recast group outperformed the model group as well. For object topicalization, however, no differences were found for any of the
three groups, nor was any learning even demonstrated. Thus, the overall findings imply that the efficacy of this type of responsive negative feedback, although evident in some cases, is questionable in others. Furthermore, some structures may be too far beyond learners’ developmental readiness to show improvement in the short term.

In another study of corrective feedback versus positive evidence alone, Muranoi (2000) found that Japanese university learners of L2 English receiving either prompts or recasts showed significantly better acquisition of indefinite articles than did learners without any type of corrective feedback, and that these results were sustained up to five weeks. Han (2002) also found that the use of recasts led to greater tense consistency in both oral and written narratives for up to five weeks. However, her study included only 8 learners and provided no results of statistical tests.

In an interesting variation on these types of comparative studies, Iwashita (2003) observed the interactions of university L2 Japanese learners in natural exchanges with L1 Japanese speakers as they worked in dyads to complete a communicative task that had been specifically designed to elicit locatives and –te forms. She found that positive evidence was much more frequent and that it led to improvement on her outcome measures for both target structures, but only for the more proficient students. Negative evidence, on the other hand,
had a beneficial effect regardless of initial learner ability.

Sanz and Morgan-Short (2004), however, have tried to claim that positive evidence and negative evidence are equally effective. In a computer-based study of 1st year university L2 Spanish learners’ interpretive and productive performance on object pronouns and word order (e.g., ‘Lo besa el niño’ The child kissed him vs. ‘Lo besa al niño’ He kissed the child) under four conditions (+/- explanation of grammatical rules before instruction, +/- explicit feedback on learner performance), the researchers found improvement for all groups with no significant differences among them. In their conclusions, they contend that their results indicate parity between positive and negative evidence. However, their experiment is largely a replication of VanPatten and Cadierno’s (1993) PI study with a slightly less proficient population. In the presentational phase, the learners saw a sentence incorporating the target structure along with pictures illustrating one correct and one incorrect interpretation, from which they had to choose. While the learners in the explicit feedback group were given a metalinguistic explanation, even learners in the supposedly ‘non-explicit’ feedback group were told “Sorry, try again” when they made the wrong choice; thus, both groups effectively received corrective feedback, albeit to varying degrees of elaboration.

A more compelling attempt to discredit negative evidence has come from Leeman
(2003), who studied 1st year university L2 Spanish learners’ acquisition of gender and number markings. As the support for the negative evidence argument in L1 acquisition all comes from studies of recasts, Leeman looked at this specific type of corrective feedback and tried to strip apart its constituent features. Consider the hypothetical student-teacher exchanges in (1) and (2) below:

(1) Student: I like dog.
Teacher: You like dogs?

(2) Teacher: You like dogs?
Student: Yes, I like.

While the teacher uses exactly the same words in both instances, only (1) qualifies as corrective feedback. Notice, however, that it not only signals that something is amiss, but also provides additional positive evidence with emphasis on the linguistic target form. Thus, the researcher sought to unravel these three factors by assigning learners to one of four conditions: recasts, error signaling without correction, normal positive evidence, and positive evidence with additional stress and a change of intonation for the target forms. She found that those in the recast and enhanced positive evidence groups outperformed the normal positive evidence group on both linguistic targets on an immediate posttest, but that only the latter difference was sustained after one week. As such, she concluded that the efficacy of recasts
can at least partly be explained by their enhancement of the salience of the target forms.

Leeman’s findings, however, are questionable on at least two grounds. First of all, as she herself states, her recasts were neither stressed nor emphasized (p.49). While it is true that context alone can heighten salience (cf. Tomlin & Villa, 1994), normal recasts also include other attention-drawing features that Leeman not only withheld during feedback but implemented in her enhanced positive evidence condition. In short, her two supposedly salient conditions do not appear to have been equally salient. The second threat to her study lies in her delivery of the enhanced positive evidence. Immediately after administering her pretest, on which her participants had had to produce the target forms themselves orally, she provided them model utterances with exaggerated stress on normally unstressed syllables. Given the fact that her participants were volunteers, who were arguably more motivated (if not higher achieving) than average learners, it is easy to imagine that they could have interpreted these artificial modifications as corrective feedback on the immediately preceding task.

Russell and Spada (2006) performed a meta-analysis of the corrective feedback research published in English between 1982 and 2004 and found an average effect size of $d = 1.04$ and a 95% confidence interval of .7 to 1.38. Using Cohen’s (1988) guidelines for
interpreting these numbers in the social sciences, where coefficients of .2, .5 and .8 can be considered small, medium, and large, respectively, corrective feedback might be seen as having at least a medium and possibly a very large effect on L2 performance, if not development. However, these figures include studies of both implicit and explicit oral and written feedback with or without prior equivalence between treatment groups or random assignment. It should also be pointed out that of 56 studies originally identified, 41 were excluded for either use of a unit of analysis other than individual participants, lack of a control group or pretest/posttest measure, or failure to provide sufficient information to calculate an effect size. Thus, the 15 studies on which Russell and Spada’s conclusions are based represent only a little over a quarter of the original body of work they examined, to say nothing of the fugitive literature, such as master’s theses, doctoral dissertations, and unpublished (usually statistically nonsignificant) studies, all of which would almost certainly reduce the magnitude of their findings.

Despite continuing debate on both sides of the argument over its necessity, the evidence suggests that corrective feedback may at least to some extent serve to accelerate the acquisition process, in which case it would be worth using. Woods (1989) opines that as long as focus on form takes place during communicative activities, it cannot do any harm, and if
learners become sensitized to the target linguistic features, so much the better. Nevertheless, to maximize any benefit, especially in view of the potential cost in terms of learner motivation, it would be useful to know specifically which types of corrective feedback are most effective. In an attempt to answer this question, the final sections of this review will examine two specific facets which have often been identified as important contributing factors, namely explicitness and salience.

2.3 The Issue of Explicitness

An issue that continues to divide even those who agree on the facilitative value of negative evidence is the extent to which learning needs to be explicit. While Schmidt (1990) contends that learner attention to linguistic forms is necessary for the conversion of input to intake, Tomlin and Villa (1994) disagree, holding that attention implies awareness, “a particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus” (p.193). They then argue that unconscious detection (i.e., cognitive registration of sensory stimuli) is possible and, therefore, the minimal requirement. In evidence, they offer the results of semantic priming studies. In a detailed examination of these and a number of other studies purporting to demonstrate implicit learning, however, Schmidt (1995) raises several problematic issues,
such as the impossibility of priming unfamiliar percepts.

Although he believes language learning to be mostly an implicit process, N. Ellis (2002b) echoes Schmidt’s position in his assessment of the importance of noticing in the initial registration of memory traces: “New associations are best learned explicitly. If they are acquired implicitly, it is only after many, many repetitions” (p.301). Specifically with regard to SLA, he also asserts that L2 learners are initially only optimally attuned to perceive stimuli relevant to their L1 and that consciousness of the importance of previously ignored stimuli creates a dialectic that can drive change in the interlanguage (N. Ellis, 2006a). These claims are consistent with Swain and Lapkin’s (1989) suggestion that it is because of greater cognitive maturity that late immersion students are able to profit from linguistic explanations and, thus, achieve L2 language proficiency comparable to that of their early immersion counterparts, despite much less exposure.

2.3.1 Implicit and Explicit Instruction

To test the role of explicit knowledge in SLA, Hulstijn and Hulstijn (1984) studied adult L2 Dutch learners who completed a story retelling task under both grammar- and information-focused conditions, in both speeded and unspeeded versions. The results showed that the grammar focus led to superior language performance even in the speeded retelling
and that learners with the highest levels of explicit knowledge performed the best.

In another investigation of implicit and explicit learning conditions, N. Ellis (1993) conducted an experiment with post-pubescent learners trying to acquire soft mutation in Welsh. As none of these learners had had any previous experience with the language, Ellis was able to observe the process untainted by prior learning. His three treatment conditions consisted of rules only (i.e., explicit teaching), exemplars only (i.e., implicit teaching), or rules plus exemplars (i.e., initial explicit teaching followed by implicit tuning). He found that learners in the implicit condition showed the fastest performance on familiar items but minimal transfer to new ones and little development of rule awareness. Also, those in the rule learning condition exhibited explicit knowledge that they often did not apply. Most interesting, however, was the Ellis’ discovery that the rules plus exemplars condition led to initially slower development but ultimately the best rule awareness and transfer to novel examples.

Another researcher who has investigated the relative efficacy of implicit versus explicit instruction has been Robinson (1996), who first did an experiment with adult ESL learners of simple and complex rules: the permissibility of subject-verb inversion in sentences fronted by adverbials of movement or location (e.g., Up the hill the mouse ran/ran
the mouse) and the formation of pseudoclefts of location (e.g., Where I want to be is anywhere but here), respectively. Thus, he sought to test Krashen’s (1985) assertions that explicit learning would only show results for the former and would, in fact, be inferior for the latter. While Krashen’s claim with respect to simple rules was supported, learners in neither of the two instructional conditions showed any differences in their development of complex rules. In another experiment with the same population, this time learning the rules of dative alternation, Robinson (1997) also found no post-treatment differences in the accuracy of grammaticality judgments of any of his learners. However, those who had received a metalinguistic explanation of the rule showed superior ability to generalize to new instances as compared to those in his implicit, incidental, or enhanced learning conditions.

The preceding experiments were all conducted in laboratories, whereas the classroom sometimes presents a different picture. For example, in a study of Canadian advanced university L2 French learners who had completed either a Core (i.e., grammar-based) or Immersion (i.e., communication-based) program in junior and senior high school, Renou (2000) found that although the former demonstrated significantly better ability to state rules on a variety of structures in a grammatical correction task, the two groups showed no differences in their abilities to make the actual corrections. Furthermore, while the Core
The group performed better on written grammaticality judgments, the Immersion group was better at oral ones. In short, the differences between the two groups could more sensibly be attributed to their respective previous training modalities (i.e., oral or written) as opposed to any benefit of explicit grammatical knowledge.

In a counter-balanced quasi-experimental design, Scott (1989) investigated the effects of implicit and explicit instruction on advanced university L2 French learners’ acquisition of relative pronouns and the form and use of the subjunctive. While she found a significant effect for explicit instruction for both structures on her written measures, she found no such difference for either in oral production. Rosa and O’Neill (1999) obtained similar results for 4th semester university L2 Spanish students learning contrary-to-fact conditionals: Those receiving explicit instruction showed greater awareness and more intake as measured by think-aloud protocols and a speeded multiple-choice recognition task, respectively.

To give a comprehensive view of the overall effects of L2 instruction, Norris and Ortega (2000) performed a meta-analysis of 49 experimental and quasi-experimental studies published between 1980 and 1998 and found a larger average effect size for explicit approaches ($d = 1.13$) over implicit ones ($d = .54$) on their targeted forms. As these results do not pertain solely to corrective feedback, however, a more focused examination, including
some more recently conducted studies is in order.

2.3.2 Implicit and Explicit Corrective Feedback

According to R. Ellis (2001), corrective feedback techniques fall along a continuum from explicit to implicit, where the former overtly states that a learner’s output is incompatible with the target language system and the latter merely intimates it. Chaudron (1977) notes, however, that implicitness or explicitness generally depends on the psychological reality of the corrective move for either the teacher or the student, which may, in fact, be different.

In giving implicit feedback, teachers most often signal that something is amiss by means of a recast, that is, a reformulation of all or part of the learner’s utterance, minus the error (Lyster & Ranta, 1997). Another implicit technique is the use of a clarification request (e.g., Excuse me?). Explicit feedback, on the other hand, is far less subtle. Although there are a number of options under this umbrella, some common ones include direct admonition (e.g., No, not …, but …), mini drills (e.g., Repeat after me), metalinguistic cues (e.g., Watch your verb), and metalinguistic explanations (e.g., Gender in English is natural, not grammatical) (cf. Ellis, Basturkmen, & Loewen, 2002).
Perhaps the best-known advocate of implicit corrective feedback is Michael Long, who endorses the use of recasts, for he claims that they are the most effective way to direct attention to form without undue detraction from an overall focus on meaningful communication (Long, 1996, 2007). In arguing for their facilitative value, he reasons that they convey necessary information about the target language in context at a moment when the learner already has a partial grasp of the semantic message and can, thus, more easily map form with function. The empirical evidence in support of Long’s position, however, is mixed.

In a study of university learners of intermediate L2 Spanish, O’Reilly, Flaitz, and Kromrey (2001) found that those who were given recasts or asked confirmation questions on oral communication tasks involving informal singular (tú), formal singular (Ud.), and formal plural (Uds.) imperative forms performed no better in their use of them after five weeks than did those who received no corrective feedback at all. Ayoun (2001), on the other hand, investigated the effects of computer-based, written recasts on 2nd, 3rd, and 4th semester university L2 French learners’ acquisition of passé composé and imparfait aspectual distinctions and concluded that they produced significantly better results in accuracy of use than did either preemptive negative evidence or positive evidence, although all three groups
improved over time. In a subsequent re-analysis of her data based on qualitative criteria, however, she later judged that it was rather the learners in the grammatical rules condition that had shown the best performance (Ayoun, 2004).

A few studies suggest that recasts may be more beneficial either for particular linguistic forms or for certain learners. Iwashita (2003), for example, observed university beginners of L2 Japanese in naturalistic exchanges with L1 users and found recasts to be the most effective type of conversational move for one of her two target structures (the bound verbal morpheme –te), but not the other (the locative-initial construction, e.g., no mae ni), which showed no differences for any treatment type. Iwashita also notes that positive evidence was more prevalent and that it led to better posttest performance on both target structures, but only for the more proficient students, whereas the benefits of recasts were unrelated to learner ability.

Philp (2003), however, in a study of adult ESL students and question forms, found learner differences of another type. Although the noticing of recasts was similar at 60 to 70%, only the more advanced learners demonstrated accurate recall. Furthermore, accuracy was additionally tied to the length of the recast and the number of changes it contained. Thus, while Iwashita’s findings suggest a possible role of salience in the efficacy of recasts, Philp’s
support the notion of learner attentional priorities on meaning as well as limitations on working memory.

Adding another piece of the puzzle is a stimulated recall experiment by Mackey, Gass, and McDonough (2000), examining ESL and Italian-as-a-Foreign-Language learners’ perceptions of corrective feedback. These researchers found that morpho-syntactic feedback most often was either misinterpreted as semantic difficulty (i.e., conversational errors) or that it went unnoticed altogether. Lexical and phonological feedback, on the other hand, were correctly perceived a much greater percentage of the time. Interesting to note, however, is the fact that morpho-syntactic errors were almost exclusively handled through simple recasts, whereas phonological and lexical problems were more often resolved through negotiation followed by recasts or through negotiation alone.

In an observational study of an elementary L2 French immersion class, however, Lyster (2001) found recasts to work better than negotiation for phonological errors but the reverse to be true for lexical and grammatical errors. One explanation for these seemingly contradictory results is that younger learners in general have more target-like pronunciation and use simpler syntactic constructions, thus making their speech more predictable, for recasts can only occur when the teacher is relatively sure of the learner’s intended utterance.
Consistent with previous studies, though, Panova and Lyster (2002) found recasts to be prevalent in the adult ESL classroom that they observed over a 10-hour period. They argue that an overabundance of recasts leave little opportunity for uptake, that is, immediate reaction to the teacher’s attempt to draw attention to the error (Lyster & Ranta, 1997). They then advocate more elicited uptake, or ‘pushed output’ (Swain, 1985), to force learners to ‘notice the gap’ (Schmidt & Frota, 1986) between the input they observe and their own usual production, which they contend will promote more learner-generated repair.

In a study of corrective feedback types employed in four primary-level French immersion classrooms, Lyster and Ranta (1997) observed six different techniques: explicit correction (e.g., *Tu veux dire la piqûre. Piqûre. Oui?* You mean the shot. Shot. Right?), clarification requests (e.g., *Pardon?* Sorry?), metalinguistic feedback (e.g., *Est-ce qu’on dit le éléphant?* Do we say a elephant?), elicitation (e.g., *Le chien peut court ? Le chien peut ...* The dog can runs? The dog can …), and repetition with rising intonation to highlight the error (e.g., *Le giraffe? The MASC-SING giraffe?*). Although recasts were the most commonly used, they were the least likely to spur to learner uptake.

In a subsequent critique, Lyster (1998) further points out that recasts may be either corrective or non-corrective and that the two types occur with near equal frequency. Thus, he
suggests that recasts may be ambiguous and perhaps insufficiently salient. Effectively, the use of a recast is much like the asking of a confirmation question, and if the learner misinterprets it as such, he or she may respond with a simple affirmative and, thus, defeat its purpose. As such, Nicholas, Lightbown, and Spada (2001) contend that recasts are most effective when their corrective nature is clear. Long (2007) counters, however, that Lyster and Ranta’s comparison techniques all result in much more uptake because they interrupt the flow of communication and oblige learners to respond to them before continuing. Moreover, Mackey and Philp (1998) note that immediate repair does not necessarily signal a change in linguistic development. Nevertheless, Lyster and Ranta argue that it is at least a sign of noticing.

Han (2002) has suggested that in order for recasts to be effective, they must focus consistently on specific forms over a short period of time in a context where the learner receives individualized attention. While admitting the practical difficulty of meeting these criteria in a classroom setting, she also suggests that it may explain the more promising results of lab-based studies.

To prevent the misinterpretation of recasts, Doughty and Varela (1998) suggest highlighting errors with a change of stress and intonation to promote noticing. Indeed, the
recasts they used in their comparison of the effects of positive vs. negative evidence in the acquisition of L2 English past forms included a special technique, wherein the teacher repeated the learners’ erroneous utterances with rising intonation in key places before making her reformulations. Thus, one could argue that the recasts themselves merely served as a model (i.e., provision of additional positive evidence) once the sources of error had already been made salient by other means. Interestingly, in a descriptive study done many years ago, Chaudron (1977) already had remarked that of the many different corrective feedback moves he observed, utterance repetition with emphasized error replacement (i.e., enhanced recasts) seemed to have the most consistent effect.

In cases of such suggested modifications, however, Ellis and Sheen (2006) question how implicit recasts really are. Perhaps realizing this problem, Long (2007) has recently issued the following re-conceptualization of the technique:

A corrective recast may be defined as a reformulation of all or part of a learner’s immediately preceding utterance in which one or more nontarget-like (lexical, grammatical, etc.) items is/are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on meaning, not language as object. (p. 77)

He then specifically adds that recasts convey corrective information implicitly and incidentally in a context where the learner’s message is at stake, thus heightening his or her
motivation and attention and consequently facilitating noticing. To better appreciate this shift of stance, compare this earlier statement from Long and Robinson (1998):

Negative feedback, including recasts, … draws learners’ attention to mismatches between input and output, that is, causes them to focus on form, and can induce noticing of the kinds of forms for which a pure diet of comprehensible input will not suffice (e.g., items that are unlearnable from positive evidence or are rare, and/or semantically lightweight, and/or perceptually nonsalient, and/or cause little or no communicative distress. (p. 23, emphasis added)

The point here is that the learner’s intended meaning is either in question or it is not, and the kinds of forms for which Long and Robinson advocate a need for negative evidence will never result in any sort of communicative breakdown: They are didactic errors. By Long’s new definition, however, corrective recasts should only occur in the context of negotiation for meaning, that is, when part of the learner’s utterance is unclear. Logically, then, that utterance would have to contain at least two errors, one conversational and one didactic. The prospects of the latter being noticed are doubtful, however, for learners have difficulty attending to meaning and form simultaneously and will almost certainly prioritize the former (VanPatten, 1990; Skehan, 1998). One of the attractions of recasts has been that (assuming they are noticed) they allow learners the choice whether to overtly acknowledge their recognition through uptake, but to follow Long’s most recent prescription would
Unfortunately erase any remaining distinction between recasts and confirmation questions.

While the eradication of conversational errors is indisputably more important than that of didactic ones, the handling of the former is also much less controversial since they tend to resolve themselves in normal social interaction. Although it would be tempting to dismiss altogether the value of accuracy in cases of didactic errors, the functions of morphological and syntactic rules are mainly formal and hardly ever semantic and may, thus, discourage learner attention (Hultijn, 1989). Furthermore, even completely redundant structures (e.g., the English third person singular indicative morpheme –s) are linguistically necessary for efficient communication as they increase processing speed on the receiving end (Shannon & Weaver, 1963). As a seemingly trivial example, when a listener hears “Do you have a …?”, to save processing time, his to her brain activates the most probable continuations in the situational context, but even on a rainy day, a speaker of standard American English would not expect the next word to be “umbrella.” In fact, that word, having already been eliminated as a possibility would require extra retrieval time and potentially disrupt subsequent processing as a result. Noting that grammatical inflections often exhibit this type of redundancy, N. Ellis (2002a) argues that they are a prime motivation for exploring the effects of explicit forms of instruction such as corrective
feedback.

In a laboratory study of Spanish-speaking adult ESL learners, Carroll and Swain (1993) asked participants in five different feedback conditions to reformulate individual sentences using dative alternation (“He built a house for us”/*“He built us a house” vs. “He constructed a house for us”/*“He constructed us a house”) after specifically training them on the process. They found that the group receiving explicit utterance rejection plus metalinguistic feedback (i.e., “Wrong” plus explanation) outperformed those receiving mere utterance rejection (i.e., “Wrong”), recasts (i.e., correct reformulations), or pragmatic disconfirmation (i.e., “Are you sure?”) but that all treatment conditions outperformed the control (i.e., no feedback). From this evidence, in addition to the superiority of metalinguistic feedback, one might conclude that explicit and implicit negative linguistic feedback are equally effective in improving L2 performance. However, the target pattern, as well as the type of corrective feedback to expect, was announced to the participants beforehand; thus, all four treatment conditions could essentially be considered explicit in this instance. Moreover, despite the fact that the entire study was conducted over a single week, the delayed posttest results indicate that performance had already started to drop.

More germane to the present study, however, is one by Muranoi (2000), who
investigated the acquisition of indefinite articles by Japanese university learners of L2 English. He found that elicitation and recasts were similarly more effective than no form-focused feedback, even after five weeks, although more so on oral than on written measures. In an interesting twist, however, he debriefed his learners differently after the treatment tasks, providing some with metalinguistic explanations and others with purely performative evaluations. Here, too, he found a significant difference in favor of the former.

Finally, in a study of the efficacy of form-focused instruction for grade 5 French immersion students learning cues for grammatical gender marking in three different feedback conditions, Lyster (2004) found mixed results. On his written tasks, up to eight weeks after the treatment, those whose output had been prompted performed better than those who had received recasts or no corrective feedback. On the oral tasks, however, no differences were revealed for any of the three groups. Important to note is that all three of these groups had undergone several weeks of initial consciousness-raising activities, including materials designed to make the target patterns more noticeable in the input. Consequently, one with an aversion toward corrective interventions might reasonably ask whether enhancing the salience of certain forms might not have the same effect. Thus, it is now to this final topic that we turn our attention.
2.4 The Importance of Perceptual Salience

As linguistic elements vary on such characteristics as phonetic substance, stress, pitch, and position, some are more readily perceived than others, that is, they have greater salience. Although they do not rule out the possibility of a role for salience in language acquisition, Dulay, Burt, and Krashen (1982) have expressed doubt, indicating that the phenomenon is at least “tempered by other, as yet unknown, factors” (p. 33). In support of their position, they offer evidence from order of acquisition studies and ask why similarly salient elements (e.g., the morpheme –s in the regular plural of *cats* vs. the third person singular indicative marker in *it hurts*) are acquired at different times or why more salient elements are sometimes acquired later than less salient ones (e.g., the long plural in *horses* vs. the short plural in *goats*). N. Ellis (2006b, 2006c) explains, however, that these phenomena can easily be accounted for from an associative learning perspective, wherein form-function mappings depend on a number of other factors as well, such as cue reliability and competition. For example, while *it hurts* ends in /s/ as a grammatical marker, the verb in *I floss* ends with the same sound; thus, verb-final /s/ is not always an indicator of grammatical meaning. Furthermore, in the case of the third person singular indicative, the more salient unbound morpheme *it* blocks the processing of the concomitant but unnecessary –s. For this reason,
Ellis argues that the linguistic forms presenting the greatest difficulty for L2 acquisition are those that “are of low salience and are redundant in the immediate understanding of an utterance” (Ellis, 2005, p. 323). In other words, some elements are relatively harder to notice (cf. Schmidt, 1990).

A common strategy for directing learners’ attention to these elusive linguistic elements is to attempt to increase their salience through a technique known as *input enhancement* (Sharwood Smith, 1991). As Sharwood Smith (1993) notes, however, the term *input* is misleading and is not synonymous with *intake* (cf. Corder, 1967), which is the processed portion; thus, he warns that input enhancement is no guarantee of consciousness-raising since the teacher’s idea of salience may not always coincide with the students’. As such, researchers have tested the effectiveness of numerous enhancement strategies, ranging from highly explicit metalinguistic description to much more implicit input flooding (i.e., increasing the relative frequency of a form’s occurrence), with the spectrum of corrective feedback moves discussed in the previous section falling in the middle (Tomlin & Villa, 1994). When operationalized in the research literature, however, most often the term *input enhancement* refers to alterations of the auditory or visual qualities of the target element itself as these types of changes are relatively unobtrusive, and therefore more palatable, to those
who prefer not to spend instructional time on rule teaching and error correction.

One notable exception to this tendency toward typographical alteration in this line of research is Williams and Evans (1998), who used an input flood to test their hypothesis that some linguistic forms might be more amenable to it than others. The researchers investigated intermediate-level ESL university students’ acquisition of participial adjectives of emotive verbs (e.g., bored vs. boring) and passives under four different learning conditions, namely flooded or unflooded input with or without negative evidence in a combination of both explicit instruction and corrective feedback. While the group receiving explicit instruction on participial adjectives showed significantly greater gains in productive accuracy than did the other two groups, the simple input flood and control groups were statistically similar. Furthermore, the group receiving explicit instruction on passives showed greater frequency of their use but no improvement in accuracy; once again, the simple input flood and control groups were not statistically different. In short, it would appear that merely increasing the frequency of a target form in the input is insufficient in promoting its acquisition. Moreover, some structures may indeed be too complex for L2 learners to acquire at certain levels of interlanguage development, even under conditions of explicit instruction.
L2 researchers who advocate the use of input enhancement frequently draw initial support from studies in L1, where there is some evidence that typographical cues lead to increased noticing. Fowler and Barker (1974), for instance, found that university students retained highlighted material significantly better than non-highlighted material, though the study lasted only one week. Hartley, Bartlett, and Branthwaite (1980) conducted a similar week-long study with 6th grade children and found that underlining not only led to better recall but, just as importantly, without cost to the non-underlined words. In a later review of this area of inquiry, however, Lorch (1989) points out that typographical cues do sometimes negatively affect unsignaled content, although there do not appear to be any differences among cue types, including underlining, boldface, color variation, italics, capitalization, and spacing variation. Furthermore, he suggests that in order to be effective, cues need to be consistent (i.e., one per information type) and their target content limited.

Even when in keeping with Lorch’s recommendations, however, applications of input enhancement to SLA have not generally produced such positive results. One exception is an early study by Doughty (1991), who conducted a computer-based investigation of university ESL learners’ acquisition of relativization under rule-oriented, meaning-oriented, and unaltered presentational conditions. In essence, the learners in all three groups were required
to read and understand a story sentence-by-sentence as it appeared on the screen. In the rule-
and meaning-oriented groups, however, the words in these sentences were animated to
illustrate their grammatical relationships. While the learners in both treatment groups
outperformed the control on oral and written measures immediately afterwards, the meaning-
oriented condition, in which the target forms were highlighted and capitalized, also led to
significantly better reading comprehension.

Another L2 researcher who has obtained positive results with input enhancement has
been Shook (1994), who studied its effects on 1st and 2nd year university L2 Spanish learners’
acquisition of relative pronouns (que/quien) and the present perfect. Using a counterbalanced
design in which the target structures were boldfaced and capitalized in one of two reading
passages, the researcher found that the enhancement improved immediate production and
recognition. More interestingly, however, he also uncovered a significantly stronger effect
for the present perfect, which he attributes to its relatively greater contribution to the
meaning of the passage. This point cannot be overemphasized, for it presents a potentially
critical difference that could explain the often conflicting results of L1 and L2 experiments.
In L1 studies, the typographically enhanced items always consist of meaningful content (i.e.,
facts) to be recalled as declarative knowledge, whereas in SLA, not only are the targets
generally semantically redundant, but they are measured in a procedural fashion to avoid the trap of the learning-acquisition argument.

In a study of 2nd semester university L2 Spanish learners acquiring preterit and imperfect verb forms, Jourdenais, Ota, Stauffer, Boyson, and Doughty (1995) also claim to have found significant positive effects for typographical enhancement. After reading a text in which the two target forms were either plain or altered, learners in the two conditions completed a writing task while following a think-aloud protocol. Statistical analyses showed that those in the enhanced condition explicitly mentioned the targeted structures a significantly greater proportion of the time. They also exhibited greater accuracy of the target forms in their written production. The credibility of these results is weakened, however, by several factors. First, the study began with only 14 participants, who were assigned to treatment groups after matching on the basis of their performance on a sub-section of a midterm examination they had taken four weeks earlier. As the scores on this instrument ranged from 2 to 10 (out of 10) with a mean of 8.14, the initial equivalency of the groups is questionable. One might also wonder why it was thought necessary to look at the chosen target structures when most of the learners had already demonstrated mastery of them by the classroom testing standard, except that no reliability information for the midterm exam was
given. Finally, the published report states that data from four learners (two from each group) had to be excluded either on account of their non-compliance with the task instructions or because of researcher error.

Confirming these suspicions, Overstreet (1998) conducted a study of 3rd semester university L2 Spanish learners acquiring the same preterit and imperfect verb forms. In a slightly more complicated design, the researcher in this case examined not only the effects of textual enhancement but of familiar versus unfamiliar content as well. His results, however, showed no positive effects for either variable on either comprehension or production. Moreover, they revealed a negative effect of enhancement on comprehension similar to those noted by Lorch (1989) above.

In a study of 2nd semester university L2 Spanish learners acquiring impersonal imperative forms, Leow (1997) compared the effects of enhanced and unenhanced texts of different lengths. While shorter texts led to better comprehension, they had no effect on intake of the target structures. Furthermore, textual enhancement showed no effects at all, positive or negative, on either of the dependent variables.

Another variable that has been compared with typographical input enhancement is preemptive negative evidence. Alanen (1995) conducted an experiment with university
students learning Finnish solely for the purpose of her investigation, in which she assigned participants to one of four treatment conditions comprising different combinations of rule instruction and textual enhancement for three locative suffixes (-ssa ‘in’, -ssa ‘on’, and –lla ‘on’) and four linguistically conditioned consonant gradations (kk→k, pp→p, tt→t, and t→d).

For the inflections, both rule conditions outperformed the control, and rules without enhanced input proved better than enhanced input without rules, but no other pairwise differences were significant. For the consonant changes, the rule-based groups simply outperformed the meaning-based groups. As the researcher had employed a think-aloud protocol, she also discovered that although most participants in the enhanced groups had indeed noticed the use of italics, not all had considered a reason for their use. Apparently, it is possible that textual input enhancement might lead to noticing the existence of a form without necessarily revealing its function.

Lending credence to this possibility is Robinson (1997), who looked at adult ESL students acquiring dative alternation of nonce forms in one of four learning conditions: implicit, incidental, enhanced, and instructed (i.e., metalinguistic). His results showed that accuracy on judging the grammaticality of the training items was the same for all groups but that the instructed group was superior at generalizing to new instances.
In another study of typographical enhancement, J. White (1998) investigated its effects on 6th grade francophone ESL learners’ acquisition of possessive determiners (e.g., *his/her*). In the first of her three treatment conditions, all instances of the target forms were enhanced in a 10-hour set of reading materials spread over two weeks. The second condition comprised not only the visually enhanced materials, but two to three hours of unenhanced enrichment reading and listening materials per week over the entire course as well. Finally, the materials in the third condition were completely unenhanced. White’s results, however, showed no statistical differences for any of the three treatment groups.

In attempting an explanation, White re-examined each of her subjects’ characteristic use of the target structure and classified it into one of eight theoretically-defined developmental stages. Although she still found no statistical differences between treatment groups, she did notice a general tendency for most learners to progress by at least one stage over the duration of the course. Moreover, when she later compared data from a fourth group of learners who had had the same teachers and curriculum but who had not taken part in the original experiment, she discovered that her three treatment groups all had a higher proportion of students in the upper stages of possessive determiner development. White then suggested that the written components of the pre- and posttests, particularly the multiple-choice section,
may have sensitized all three of her treatment groups to the peculiar nature of the English
target structure vis-à-vis its dissimilar French counterpart. Furthermore, she conjectured that
the mismatch between L1 and L2 gender agreement rules (i.e., natural vs. grammatical) may
have been too complex for these learners to grasp without much more explicit attention being
drawn to the relationships between the determiners and their referents, either through direct
instruction, corrective feedback, task explanation, or some other method of input enhancement,
such as color coding. In fact, in a follow-up questionnaire in which she asked the learners
why they thought some of the words had been altered, only one-third of her respondents
named the target forms.

Leow (2001) then asked not whether input enhancement necessarily leads to

\textit{acquisition} but whether it can even promote greater \textit{noticing}. In a study involving the

formal/polite (i.e., \textit{usted}) form of the imperative, which had not been introduced in the 52

hours of instruction that the students had already completed, 1\textsuperscript{st} year university L2 Spanish

learners were randomly assigned reading materials in either an enhanced or an unenhanced
treatment condition, in the former of which all the target verb forms were underlined with
their endings highlighted in boldface. In order to observe their noticing, the researcher had
the learners verbalize their thoughts as they circled any unfamiliar verb forms they
encountered. Three weeks after the treatment, however, his results showed no differences in noticing or comprehension between the two groups. The effects of noticing on immediate intake also showed to be nonsignificant.

By way of explanation, Leow suggests that his instructions for his subjects to mark any unfamiliar verb forms may have been sufficient to promote their noticing, for although no significant difference was found between the two groups, both showed high levels of sensitivity to the target structure in their think-aloud protocols. Another point of interest was that some students in the enhancement group not only mentioned the typographical cues but expressed puzzlement over them as well. The most intriguing aspect of this study, however, was the performance of a pair of outliers, one in each group, who obtained near perfect scores on all measures, including the delayed written production task, and who explicitly mentioned in their think-aloud protocols their musings on the significance of the unfamiliar forms.

Izumi (2002) found similarly disappointing results for input enhancement in a comparison with pushed output for university ESL learners studying relativization. Although in his case enhanced input did lead to significantly more noticing, it was pushed output that led to greater accuracy of use. Thus, once again, simple awareness of form did not lead to
improved control.

One final study that deserves mention, however, is that of Leeman (2003), whose findings not only contradict the majority of all others but reopen the case against negative evidence. In an experimental design involving 1st year university L2 Spanish students in an oral communication task, Leeman not only managed to control the input to her four treatment groups but attempted to isolate target structure salience and corrective feedback type as well. She concluded that the former, rather than the latter, was the critical factor. Leeman’s enhanced salience condition, however, could easily have been interpreted as corrective feedback as it obviously targeted the same linguistic structures (i.e., adjective and number agreement) as the immediately preceding pretest, a production task at which the participants had clearly failed. Finally, she conducted her study on volunteers, who are already arguably predisposed better to interpret teacher feedback the way it is intended.

In short, input enhancement, at least in its most common forms, does not seem to be particularly effective at promoting noticing and, even if it were, there is no indication that mere sensitization actually leads to intake. Perhaps metacognitive awareness, as was evident with Leow’s outliers, is a requisite condition of interlanguage change. Nevertheless, in addition to carefully investigating the effects of variously explicit levels of
corrective feedback, especially in light of Leeman’s challenge, we need to reconsider the possibility that positive evidence alone, if enhanced in salience, might be sufficient for promoting the acquisition of L2 linguistic forms bearing light communicative loads.

2.5 Summary

While the verdict is still out on whether corrective feedback really works or not, as far back as 1977, Chaudron indicated that the most effective type of teacher feedback move appeared to be partial repetition of learners’ utterances with emphasis on their errors. In light of the preceding literature review, his recommendation certainly makes sense, for this type of feedback simultaneously accomplishes three essential objectives: making errors salient, reducing the burden on working memory, and pushing learner output.

Although some linguistic features, especially those that are not shared by the learners’ L1 and that do not cause communicative breakdown in normal interaction, may be difficult to notice, it is unclear whether they need to be brought to attention through corrective feedback or whether less obtrusive measures are sufficient. Furthermore, it remains to be established whether conscious knowledge of rules profits learners in integrated language use over the long term.
As much of the previous research has produced mixed results, the present study aims to provide clarity on the issues of input enhancement and corrective feedback by systematically examining them in a tightly controlled computer-based experiment, which affords numerous possibilities that essentially remain closed to the traditional classroom-based alternative. For instance, while L2 learners sometimes incidentally notice certain aspects of the linguistic code, they do not seem to develop and test their own hypotheses without pushed output and corrective feedback; thus, whether a traditional written text is typographically enhanced or not may make little difference to most learners, who may be unwilling to expend the considerable amount of time and energy needed to privately interact with it. Similar methods of enhancement with a computer, on the other hand, may at first glance appear no different, but the interactive medium can also engage learners and perhaps better maintain their attention. Consequently, learner efforts would be sustained over long periods of time, and eventually the target structures could become automatized.

Another advantage of computers is that they can efficiently collect data. Although think-aloud protocols may still have their usefulness, in some ways they are unnatural, which is why they necessitate learner training before they can be successfully implemented. Computers, on the other hand, can monitor learners’ thoughts indirectly by registering
keystrokes and mouse clicks. While this approach may have its own shortcomings, it is arguably much less obtrusive as it allows the learner not only to remain focused on the immediate task but to do so with minimal linguistic code switching.

In short, the present study hopes to shed additional light on the combinations of input enhancement and corrective feedback type that are most effective at raising learner awareness and thus facilitating acquisition of the target forms. The following chapter will describe in detail the exact methodological procedures that were employed in the investigation.
CHAPTER 3

METHODOLOGY

As the literature review in the preceding chapter revealed, despite a sizable number of studies on the benefits of formal instruction over natural exposure, the need for negative evidence, the degree to which correction needs to be explicit, and the effectiveness of input enhancement in promoting learner noticing of linguistic forms, so far there has been little irrefutable support for the general efficacy of either corrective feedback or target form enhancement in promoting sustained growth in grammatical accuracy. Thus, the purpose of the present study was to improve on some of the previous research designs and further probe this issue by investigating the effects of four types of corrective feedback (meaning-focused, implicit form-focused, non-metalinguistic explicit form-focused, and metalinguistic) and two levels of target form enhancement (unenhanced and enhanced).

This chapter begins with a description of the study design, a statement of the research questions, and an explanation of the hypotheses. It then characterizes the study participants and the linguistic target structures before detailing the outcome measures and instructional treatment materials. Finally, it relates the general study conduct as well as the data collection and handling procedures.
3.1 Study Design

The study comprised a 4 x 2 pretest-posttest-delayed posttest mixed experimental design, with both between-subjects and within-subjects components. The treatment factors were input enhancement and corrective feedback type. Input enhancement consisted of the presence or absence of typographical alterations of the linguistic target structures (i.e., à, au, en, and aux before geographic place names), during either their initial presentation or the provision of corrective feedback, the four levels of which were defined as follows:

**Meaning-focused (FoM):** acceptance of grammatically inaccurate learner utterances as long as they are informationally correct (see p. 149, Fig. 13, for example)

**Implicit Form-focused:** tacit disapproval of grammatically inaccurate learner utterances via recasts (see p. 154, Fig. 16, for example)

**Non-metalinguistic Explicit Form-focused:** unequivocal rejection of grammatically inaccurate learner utterances (see p. 156, Fig. 18, for example)

**Metalinguistic:** unequivocal rejection of grammatically inaccurate learner utterances, followed by the presentation of an appropriate linguistic rule (see p. 158, Fig. 20, for example)
The combination of these two variables yielded eight possible treatment conditions (see Figure 1 below).

<table>
<thead>
<tr>
<th>Input Enhancement</th>
<th>Corrective Feedback Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FoM</td>
</tr>
<tr>
<td>- Enhanced</td>
<td>Condition 1</td>
</tr>
<tr>
<td>+ Enhanced</td>
<td>Condition 2</td>
</tr>
</tbody>
</table>

FIGURE 1. Treatment conditions by corrective feedback type and input enhancement level.

The dependent variable was grammatical accuracy, as measured by raw numbers of correctly supplied target prepositions on one of two cloze tests that were specially designed for use in this study.

Each treatment condition informed a unique version of a computerized series of narratives with follow-up questions to check learner comprehension as well as provide opportunities for production of the target linguistic structures. To compensate for its inability to evaluate learner speech, the computer simulated oral interaction in a written mode, namely though the provision of immediate feedback. Moreover, it afforded the advantage of strictly
controlling the manner and amount of both input and correction. While lab-based studies may have limited generalizability as compared to their classroom counterparts, they also engender numerous benefits in terms of their internal validity and statistical power, such as the possibility of random assignment to treatment conditions within intact groups and the avoidance of the unit of analysis problem (cf. Levin & O’Donnell, 1999), not to mention the facilitation of simultaneous data collection from multiple participants. In short, any real treatment effect for target form enhancement and/or corrective feedback is more likely to appear in a lab setting, and once its existence has been established, we can begin to attempt its replication in the classroom.

3.2 Research Questions

The inquiry focused on the following three questions:

Research Question #1: Do L2 French learners experiencing special instructional materials incorporating the prepositions à, au, en, and aux before geographic place names show improved accuracy on these target forms over time?

Research Question #2: Are there immediate differences in target form accuracy among learners receiving different types of corrective feedback and levels of typographical input enhancement?
Research Question #3: Are there long-term differences in target form accuracy among learners receiving different levels of these same two variables?

3.3 Hypotheses

Given the penury of previous attempts to measure long-term effects in this area of research and the lack of significant findings in the few that have, hypotheses for the present study were formulated only with respect to immediate post-treatment results (i.e., for Research Questions #1 and #2). Below are the predictions, along with an explanation of each.

With respect to the first research question, perhaps obviously, it was predicted that learners across all conditions would show improved accuracy on the target structures at least shortly after the treatment lesson, for the remaining inquiries would otherwise be irrelevant. Ideally, these performance gains would also be sustained, if not increased, over the entire course of the project if they are truly indicative of bona fide acquisition as opposed to mere monitored learning. However, as the present study is the first one known on record to investigate the specific target structures in question, there was no existing evidence on which to base such a hypothesis that any change in linguistic behavior would be durable.

As for the second research question, the picture is considerably more complicated. While target form salience may promote noticing, its role in inducing the necessary cognitive
processing to trigger changes in learners’ interlanguage is less certain. Thus, it was hypothesized that the conditions that would promote the greatest accuracy on the target forms would be those that not only highlighted them but which first emphasized their importance, effectively pressuring the learners to use them correctly. As such, it was predicted that learners receiving an unequivocal signal of a need to change their linguistic behavior (i.e., explicit corrective feedback) would perform better than those receiving no such signal (i.e., either implicit or meaning-focused corrective feedback). Also expected within the explicit corrective feedback condition was an advantage for input enhancement as it would draw attention to the source of error. Given the limited duration of the proposed treatment and the communicative nature of the intended instructional materials, however, no further differences were predicted for rule explanation. In short, the lesson would not provide enough controlled practice to develop automaticity, and the accuracy measures were purposely designed to de-emphasize conscious rule application.

After the learners in the explicit correction conditions, the next highest performing group was predicted to be those receiving enhanced recasts. The reasoning was that the additional prominence would highlight the discrepancy between the positive affective feedback of communicative approval and the negative cognitive feedback of the grammatical
reformulation (cf. Vigil & Oller, 1976), which some learners may otherwise fail to notice or more easily choose to ignore. Here, the potential influence of motivation in learners’ decisions over how to use this feedback was initially considered. Unfortunately, a much larger number of participants than that which was available would have been necessary to detect such a three-way interaction (i.e., feedback type x target form enhancement x motivation level). Thus, it was merely expected that the effects in this condition would be attenuated as compared with those in the explicit corrective feedback condition.

In the unenhanced recast condition, it was predicted that the learners would not even notice the conflicting aspect of the cognitive feedback. Similarly, the learners exposed to the enhanced input but given no form-focused correction may notice the forms but see no reason to cognitively process them, and those neither exposed to enhanced input nor given any form-focused correction would overlook the target structures altogether. In short, learners in none of these three groups would have any impetus to change their current linguistic behavior.

To recapitulate, the hypotheses for the immediate posttest results were as follows:

**Hypothesis #1:** Learners experiencing special instructional materials incorporating the prepositions à, au, en, and aux before geographic place names will show improved accuracy
on these target forms afterward.

**Hypothesis #2a**: Learners receiving explicit corrective feedback (with or without rule explanation) plus enhancement of the target forms (i.e., Conditions 6 & 8) will exhibit the greatest accuracy on immediate post-treatment measures of those forms.

**Hypothesis #2b**: Learners receiving explicit corrective feedback (with or without rule explanation) but unenhanced target forms (i.e., Conditions 5 & 7) will exhibit less accuracy on immediate post-treatment measures of those forms than will their counterparts receiving input enhancement (i.e., Conditions 6 & 8) but greater accuracy than those learners in all other groups (i.e., Conditions 1-4).

**Hypothesis #2c**: Learners receiving implicit corrective feedback plus enhancement of the target forms (i.e., Condition 4) will exhibit less accuracy on immediate post-treatment measures of those forms than will their counterparts receiving explicit corrective feedback (with or without input enhancement, i.e., Conditions 5-8) but greater accuracy than those learners in all other groups (i.e., Conditions 1-3).

**Hypothesis #2d**: Learners receiving implicit corrective feedback but unenhanced target forms (i.e., Condition 3) and those receiving no form-focused correction (with or without input enhancement, i.e., Conditions 1-2) will all similarly show less accuracy on immediate post-
treatment measures of those forms than will all other groups (i.e., Conditions 4-8).

Thus, in general, it was expected that participants across all conditions would show greater accuracy in their use of the target prepositions after the treatment, at least in the short run. Furthermore, explicit form-focused feedback would lead to more accurate performance than implicit form-focused feedback, which in turn would be superior to meaning-focused feedback. Finally, typographically enhanced target forms would promote better results than would their unenhanced counterparts, but the magnitude of this effect would vary according to the type of corrective feedback involved.

3.4 Participants

The participants were 136 fluent English-speaking students 18 years of age or older enrolled in 2nd semester beginning French at the University of Arizona throughout the fall of 2006. According to self-report data collected at the time of the treatment, 82.5% of these students were between the ages of 18 and 20 by the end of the study, and most were taking the course as a general education requirement (see Tables 1 and 2 below).
TABLE 1. Student Self-reported Age at End of Course

<table>
<thead>
<tr>
<th>Age</th>
<th>Respondents</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>18 to 20</td>
<td>112</td>
<td>82.4</td>
<td></td>
</tr>
<tr>
<td>21 to 23</td>
<td>15</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>Over 23</td>
<td>5</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>Did Not Complete Survey</td>
<td>4</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>136</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2. Student Self-reported Reasons for Enrolling in Course

<table>
<thead>
<tr>
<th>Reason</th>
<th>Respondents</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>General Education Requirement</td>
<td>87</td>
<td>64.0</td>
<td></td>
</tr>
<tr>
<td>Part of Major</td>
<td>17</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>Part of Minor</td>
<td>12</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Elective</td>
<td>13</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Did Not Complete Survey</td>
<td>4</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>136</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>
As this research spanned over an entire semester, the course level was chosen to ensure that the learners would have adequate French proficiency to participate from the initial stages of the experiment but minimal previous exposure to the linguistic target structures, which were part of the official French 102 curriculum. To minimize participant attrition over sessions as well as avoid reliance on volunteers, who may not be representative of the general population, all aspects of the investigation were conducted during regular class meetings as part of the students’ normal lessons. However, to assure the students that their decision whether to allow the use of their data for the purposes of this project would not affect their grades, human subjects recruiting was conducted at the end of the semester, when performance evaluations on all relevant activities had already been communicated.

The potential participant pool originally comprised 221 learners in 12 course sections. However, all 21 members in one section were excluded because of a timing conflict that prevented them from following the same implementation schedule. Also disqualified were 46 students who failed to complete one or more of the study components: 1) a pre-treatment assessment of target structure accuracy, 2) a set of instructional materials specially designed to teach and/or reinforce the target structures, 3) an immediate post-treatment measure of target structure accuracy, or 4) a delayed post-treatment measure of target structure accuracy.
One other student achieved a score of over 80% on the pretest and was considered ineligible on grounds of having already mastered the target structures. Finally, two additional students were under the age of 18 and, thus, unable to participate without parental consent. Thus, not accounting for absences on the day of recruitment, 136 of the remaining 151 students (90.07%) granted permission to use their data for this research. Although these exclusions took place after the data collection, none was ostensibly treatment-related. As such, attrition is assumed to have been random across treatment conditions.

3.5 Linguistic Target Structures

Unlike English, French makes no distinction between static and dynamic spatial relations in its use of prepositions before complements of place. Thus, the French à can mean either ‘to’ (1a), ‘at,’ or ‘in’ (1b).

(1) a. *Je vais à la bibliothèque.*
   ‘I’m going to the library.’

   b. *J’étudie à la bibliothèque.*
   ‘I study at/in the library.’

One complication, however, is that the form of this preposition changes depending on the definite article that precedes the place name in its nominative case form (Booth, 2003). A
highly frequent structure, it begins to appear in expressions with common place names very early in formal classroom instruction (e.g., the first chapter of Débuts: An introduction to French, the two-semester adopted textbook for the elementary course at the University of Arizona):

*Mais qui est-ce à la table?*  ‘But who is that at the table?’  
(Siskin, Williams-Gascon, & Field, 2003, p.11)

*Elle est à l’école avec des amis.* ‘She’s at school with some friends.’  
(ibid, p.21)

In fact, these initial occurrences take place even before the introduction of à as a separate lexical item (p.38), when the English equivalent is unfortunately given as ‘to’ or ‘at,’ whereas idiomatically it often translates as ‘in.’ In any case, the rules for its morphology are taught soon afterward (i.e., in Chapter 3):

Wherever à is used with le and les, it forms a contraction. No contraction is made with la or l’.

\[
\begin{align*}
à + \text{le} & \rightarrow \text{au} \\
à + \text{la} & \rightarrow \text{à la} \\
à + \text{l’} & \rightarrow \text{à l’} \\
à + \text{les} & \rightarrow \text{aux}
\end{align*}
\]

(Siskin, Williams-Gascon, & Field, 2003, p. 72, boldfacing in the original)
As à has multiple functions (e.g., not only preposition of position and motion but also indirect object marker) and the textbook does not limit its examples to the one of interest, a more relevant set is proposed in (2) below:

(2) Variants of à before common place names

- à + le restaurant ‘the restaurant’ → au restaurant ‘at/to the restaurant’
- à + la banque ‘the bank’ → à la banque ‘at/to the bank’
- à + l’aéroport ‘to the airport’ → à l’aéroport ‘at/to the airport’
- à + les toilettes ‘the restroom’ → aux toilettes ‘in/to the restroom’

The adopted coursebook provides equally early incidental exposure to the similar but slightly more complicated transformations of this preposition and their distribution before geographic place, first at the end of a nursery rhyme and again in an illustration of the idiomatic use of the impersonal pronoun on ‘one’:

1, 2, 3, j’irai dans les bois
4, 5, 6, cuillir des cerises
7, 8, 9, dans un panier neuf.
10, 11, 12, elles seront toutes rouges,
    à Toulouse.

En France, on aime le pain.

(Siskin, Williams-Gascon, & Field, 2003, p.6)

In France, people like bread.
(ibid, p.27)

However, the rules are not formally presented until much later (i.e., in the fourth
chapter of the second semester):

To express the idea of in, at, or to, the preposition depends on the gender of the place name that follows it.

1. For cities, use à.

   *Hélène habite à Montréal.* ‘Hélène lives in Montreal.’
   *Camille habite à Paris.* ‘Camille lives in Paris.’

2. For continents, and for countries, states, and provinces that are feminine or that start with a vowel sound, use en.

   *Hélène arrive en Europe en avion.* ‘Hélène is arriving in Europe by plane.’
   *Elle fait un reportage en France.* ‘She is doing a story in France.’
   *Le reporter qui vient de Jerusalem rentre en Israël.* ‘The reporter from Jerusalem is returning home to Israel.’

3. For masculine countries, states, or provinces, and for plural countries, use à and form a contraction with the article.

   *Hélène habite au Canada.* ‘Hélène lives in Canada.’
   *Elle n’habite pas aux États-Unis.* ‘She doesn’t live in the United States.’

(Siskin, Williams-Gascon, & Field, 2003, p. 319, translations mine, boldfacing in the original)

Incidentally, the textbook’s initial statement in this case is misleading, for the uses of à and en are conditioned by other factors as well. While the alternate set of examples proposed in (3) below does not account for exceptional cases such as Israël, it at least
maintains consistency with those in (2) and hopefully provides an additional modicum of clarity:

(3) Variants of à before geographic place names

\[
\begin{align*}
\text{à + (no article) } & \text{ Paris ‘Paris’} & \to & \text{ à Paris ‘in/to Paris’} \\
\text{à + le Canada ‘Canada’} & \to & \text{ au Canada ‘in/to Canada’} \\
\text{à + la France ‘France’} & \to & \text{ en France ‘in/to France’} \\
\text{à + l’Algérie ‘Algeria’} & \to & \text{ en Algérie ‘in/to Algeria’} \\
\text{à + les Etats-Unis ‘the US’} & \to & \text{ aux Etats-Unis ‘in/to the US’}
\end{align*}
\]

The purpose here is not to expose the weaknesses in the way these rules are presented in the adopted coursebook but to point out that they are explicitly taught as part of the official curriculum. Moreover, evidence from written work collected in prior semesters shows that most students do not use these structures independently by the end of 1st year French. In fact, informal personal observations of the Table Francophone, a weekly extracurricular event at which students, instructors, and other French-speaking members of the community meet to converse informally, have further revealed that even many motivated late 2nd year students regularly misuse the target forms.

Besides the apparent difficulty learners have in mastering the use of these structures, another important criterion in their selection for this study was that they constitute a source of didactic, as opposed to conversational, error. In other words, a response such as the one
illustrated in the student-teacher exchange shown below in (4) would be disingenuous as the meaning of the original statement is perfectly clear.

(4) Student-teacher exchange illustrating negotiation of a didactic error

Student:  *L’été dernier je suis allée en Canada.
         ‘Last summer I went in Canada.’
Teacher :  Au Canada?
         ‘To Canada?’

Finally, the à/au/en/aux distinction before geographical place names was chosen because it is part of the official French 102 curriculum. The fact that it would not be explicitly taught by the time of the intervention would justify the integration of this research project into regularly scheduled classroom activities.

3.6 Outcome Measures

As was pointed out in Chapter 2, one difficulty in resolving the SLA debate over the efficacy of anything other than comprehensible input has been the reliance of previous studies on either de-contextualized, discrete-point measures or metalinguistic tasks, both of which draw on explicit knowledge and, thus, fail to tap acquisition as opposed to learning. Thus, in an effort to gauge not immediate conscious learning but long-term development and true communicative competence, this study employed a cloze test, an integrative measure that
requires simultaneous processing of multiple language components as does real communication (Madsen, 1983). Although scholars such as Ayoun (2001, 2004) continue to utilize them in their research, cloze tests seem to generate much less attention today than in previous decades. As such, below is a brief explanation of the cloze procedure and a rationale for its use.

3.6.1 General Cloze Procedure

In its most basic form, a cloze consists of a written text that begins and ends with one or two intact sentences to orient the reader to its general content but then contains blanks throughout, where individual words have been deleted, traditionally at regularly spaced intervals (e.g., every 5\textsuperscript{th}, 6\textsuperscript{th}, or 7\textsuperscript{th} word) everywhere in between (Cohen, 1980). The reader’s task is to then fill in each blank with a single word that both fits grammatically and makes sense. On the surface, the cloze resembles a typical sentence completion exercise. However, as the positioning of its constituent blanks is systematically governed, this type of test generally focuses not on discrete lexical or grammatical items but on global language proficiency.

For the purposes of this study, however, it was not only necessary to devise an integrative measure of the learners’ language proficiency but to reliably tap their developing
competence with the specifically targeted linguistic structures as well. Cohen (1980) rightly warns that the deliberate deletion of selected grammar points can undesirably turn a cloze test into a discrete-point measure, essentially defeating its purpose. To use the cloze method to measure grammar achievement, he suggests creating a passage that contains the relevant target structures and then deleting every \( n \)th word. Thus, this proviso was kept in mind during the development of the relevant reading materials.

### 3.6.2 Reading Passage Development

As the envisioned research design entailed repeated measures, to minimize testing effects it was necessary to create two different reading passages. The first of these dealt with the French-speaking world (‘Le Monde francophone’), the second with the Muslim world (‘Le Monde musulman’). These topics were chosen so as to be thematically similar not only to each other but also to the students’ other course materials.

In turn, these selections later helped in part to determine the content of the treatment lesson, for the plan was to include exactly eight non-overlapping geographic place names from each text (two for each of the four target prepositions). To eventually test the learners’ ability to apply their linguistic knowledge to unfamiliar items, each reading passage featured 16 additional place names (four for each of the target prepositions) as well. With geographic
regions as the organizing principle, a combination of 24 countries and cities was devised to meet the grammatical criteria for each of the two passages. The references for the content of these two texts included a variety of print and online sources, most notably Walter’s (1988) *Le Français dans tous les sens*, the official website of La Francophonie (http://www.francophonie.org), Wikipedia, and Islam 101.com.

Both passages underwent multiple drafts, each of which was meticulously reviewed and edited by a native-speaking French professor. The final products were nearly equal in length at 495 and 482 words.

### 3.6.3 Actual Cloze Construction

Once the reading passages were complete, the next step was to transform them into cloze exercises. Here, however, the approach adopted in this study deviated slightly from the one suggested by Cohen (1980), for Hughes (2003) recommends discretionary judgment in selecting words for omission, pointing out that a strict n<sup>th</sup> word deletion pattern nearly always produces problematic items. In practice, a two-step process was followed. First, the 24 target prepositions in each passage were deleted. Then, other words were deleted to approximate a pattern of every 7<sup>th</sup> word. For instance, if 11 words came between two target prepositions, the 6<sup>th</sup> was deleted, except in cases of proper nouns, numbers, or uncommon lexical items with
few synonyms. In the end, each passage contained 60 blanks corresponding to 24 targets and
36 non-targets and appearing every 5th to 9th word, with every 7th word being the most
frequent. (See Appendices E and F for copies of the tests.)

Originally, the cloze tests were tried on five of the six teachers (other than the author)
who were scheduled to teach French 102 during the period of the data collection. (The 6th
teacher unfortunately did not respond to any of several requests for her participation.)

Surprisingly, however, even among the five teachers who cooperated, three of whom were
educated in French through secondary school, none demonstrated perfect mastery of the
target structures in accordance with prescriptive rules. In fact, the average number of
incorrect responses on the target items was 2.4 on each form, and all five teachers made at
least one error on both. The tests were then administered to five L1 French-speaking
language teachers, none of whom gave any incorrect target prepositions on either form.

Incidentally, even the native speakers did not supply the sought-after prepositions for
all 24 target items. In fact, the scores ranged from 22 to 24 (mean = 22.8) for the text on the
French-speaking world and from 23 to 24 (mean = 23.6) for the one on the Muslim world.

While every effort was made to force production of the target prepositions in all 24 instances,
in at least one case not only the locative but also the genitive was acceptable, as in illustrated
in (5) and (6), respectively.

(5) *la religion majoritaire aux Maldives*
    ‘the majority religion *in* the Maldives’

(6) *la religion majoritaire des Maldives*
    ‘the majority religion *of* the Maldives’

Unfortunately, reformulating the sentence to disallow the latter possibility would have complicated the syntax considerably, so it was left unmodified. More often than an acceptable alternate response, however, the test taker simply put an erroneous word, but Cohen (1980) notes that even native speakers may not always score 100% on this type of task because of the concentration it demands.

3.7 Instructional Treatment Materials

Chief among the difficulties in drawing clear conclusions from previous studies of corrective feedback has been their failure to control the target linguistic input. Thus, to ensure consistent delivery of the former as well as exposure to the latter, the present investigation drew on the work of Ayoun (2001) to create a computer-based, interactive reading lesson, only in Macromedia Director rather than HyperCard. In essence, the aim was to simulate classroom conversational exchanges through multiple cycles of informational
input, elicited output, and immediate feedback. Input enhancement, then, would consist of typographically altered forms, which Doughty (1988) characterizes as the visual equivalent of stress and emphasis in aural input.

### 3.7.1 Basic Setup

While each treatment condition required its own unique version of the program, the organization and content in all instances was the same. Entitled «Le Tour du monde en 100 jours» ‘Around the World in 100 Days,’ the lesson centered on a series of 14 postcards written by a fictional French university student named Patrice, who was supposed to be traveling around the world on a semester-at-sea program. Although his story was made to resemble a real semester-at-sea experience, for methodological reasons his itinerary was dictated in part by the additional constraint of having to include four locations requiring each of the four different prepositions under investigation (i.e., à, au, en, and aux) for a total of 16 places, half of which figured in each of the two linguistic outcome measures. The maps in Comeau, Lamoureux, and Tanvouez’s (1999) *Culture et Société (6th ed.*) served as references in weighing the various possible combinations and deciding the final travel route. Figure 2 below shows a chronological list of these places, along with an indication of their associated prepositions.
<table>
<thead>
<tr>
<th>Toponym</th>
<th>Associated preposition</th>
<th>à</th>
<th>au</th>
<th>en</th>
<th>aux</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>les États-Unis</em></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>‘the United States’</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>New York</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘New York’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>les Petites Antilles</em></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>‘the Lesser Antilles’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>les îles Marquises</em></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>‘the Marquesas Islands’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tahiti</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Tahiti’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>les Philippines</em></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>‘the Philippines’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>le Viêt-Nam</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Viet-Nam’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>l’Indonésie</em></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>‘Indonesia’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>l’Inde</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘India’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pondichéry</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Pondicherry’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>le Pakistan</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Pakistan’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Madagascar</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Madagascar’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>l’Afrique du Sud</em></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘South Africa’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>le Cameroun</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Cameroon’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>le Maroc</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Morocco’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>la France</em></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘France’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

FIGURE 2. List of instructionally presented toponyms and their associated prepositions in chronological order of appearance in the lesson.
Beginning with that which he sends from within France on the eve of his departure, each postcard recounts Patrice’s recent activities at his present location and tells where he is headed next, modeling the target preposition in the process. Thus, the first postcard makes reference to his embarkation for the United States and the last announces his return to France (see Figure 3 below for an example of the former).

FIGURE 3. Typical reading passage: Postcard 1 (from France), modeling aux Etats-Unis.
After each reading passage, the lesson branches to a series of at least three related questions: generally, one short answer, one true/false, and one multiple choice (see Figures 4-6, respectively). The short answer question, as in (7) below, aims to elicit learner production of a geographical place name requiring one of the target prepositions.

(7)  *Où va Patrice pour la première étape de son voyage?*

‘Where is Patrice headed on the first leg of his journey?’

The true/false (8) and multiple choice (9) questions, on the other hand, check comprehension of other details.

(8)  *Patrice promet à sa copine de lui écrire tous les jours pendant son voyage.*

‘Patrice promises to write his girlfriend every day of his trip.’

- *Vrai* ‘True’
- *Faux* ‘False’

(9)  *Qu’est-ce que Patrice va faire au moment où il termine son message?*

‘What is Patrice going to do after he finishes his message?’

- A. *Dîner.* ‘Eat dinner.’
- B. *Dormir.* ‘Go to bed.’
- C. *Faire sa valise.* ‘Pack his bags.’
FIGURE 4. Short answer question: Part 1 of 3 in follow-up sequence to Postcard 1 (from France), designed to elicit the response *aux Etats-Unis.*
FIGURE 5. True/false question: Part 2 of 3 in follow-up sequence for Postcard 1 (from France).
The order of these formats varies slightly between passages to match that of the appearance of the answers to their corresponding questions in the reading passage. However, as the program was designed to control exposure to the target linguistic input, it does not allow students to return to the reading passage once they have proceeded to the questions. In this sense, the learning activity is arguably memory-dependent and somewhat rote, but such is unfortunately the case of most closed-ended questioning, whether computer-based or otherwise. Furthermore, one could also contend that requiring recall of details without further
reference to the text in this manner not only introduces an additional element of engagement but encourages attempts at integrated meaning-making by discouraging skimming.

While most postcards model only a single target preposition (for the next country on the itinerary), those from the United States (see Figure 7 below) and India contain an additional reference to the actual port of call (i.e., New York and Pondicherry, respectively), which explains how there were 16 target locations but only 14 postcards.

The reason for incorporating two targets into a single postcard in these instances rather than presenting them separately was to ensure elicitation of both a city and a country name with a short answer question while maintaining consistency with the pattern of one postcard per country. For example, if Patrice had mentioned in his initial postcard his intention to visit New York, the student might have answered either à New York ‘to New York’ or aux Etats-Unis ‘to the United States’ in response to the question of where Patrice was headed next. To distinguish these two possibilities and elicit them in logical sequence, another option would have been to ask to which country and then to which city he was going. However, it was felt that using the French word for to at the head of the question might adversely influence the learners’ subsequent production of it, especially since the equivalent in this case is neither à, au, aux, nor en, but dans, conditioned by the indefinite determiner quel(le) (‘which’) (Vandeloise, 1986/1991), as illustrated in (10) below.

(10) a. To which country is Patrice headed next?
   ‘Dans quel pays va Patrice ensuite ?’

b. To which city is he headed?
   ‘Dans quelle ville va-t-il?’

Thus, in the two cases where the target preposition preceded a city name, the relevant postcard also modeled a second preposition for the next country (e.g., aux Petites Antilles ‘to
the Lesser Antilles’). Moreover, the sequence of follow-up questions included two in short answer format (a and c) with one other response type (c) in between, such as in (11).

(11) a. *Où est Patrice aux Etats-Unis?*  
‘Where in the US is Patrice?’

b. *Pendant son séjour à New York, Patrice est allé voir une comédie musicale à Broadway.*  
‘During his stay in New York, Patrice went to see a Broadway musical.’

\[\begin{array}{ll} 
Vrai & \text{‘True’} \\
Faux & \text{‘False’} 
\end{array}\]

c. *Après son séjour à New York, où va Patrice?*  
‘Where is Patrice headed after his stay in New York?’

### 3.7.2 Feedback Commonalities Across Treatment Groups

Just as the narrative and follow-up questions were the same for all students, so were the programmed feedback options to the true/false and multiple choice items as they were not the focus of this study. As shown in (12), for the true/false questions, the computer either affirmed or rejected the student’s answer and then provided the same one-sentence elaboration.

(12) Computer-student interactions for correct and incorrect true/false questions

a. Computer: *Patrice promet à sa copine de lui écrire tous les jours pendant son voyage.*
‘Patrice promises to write his girlfriend every day of his trip.’

Student:  
*Faux.* ‘False.’

Computer:  
*Oui, c’est faux.* ‘Yes, that’s false.’

*Il a promis de lui écrire de chaque pays, pas chaque jour.*

‘He promised to write her from every country, not every day.’

b.   

Computer:  
*Patrice promet à sa copine de lui écrire tous les jours pendant son voyage.*

‘Patrice promises to write his girlfriend every day of his trip.’

Student:  
*Vrai.* ‘True.’

Computer:  
*Non, c’est faux.* ‘No, that’s false.’

*Il a promis de lui écrire de chaque pays, pas chaque jour.*

‘He promised to write her from every country, not every day.’

For the multiple choice questions, it either affirmed the student’s selection or rejected it and then gave the correct answer as part of a complete sentence, as in (13) below.

(13) Computer-student interactions for correct and incorrect multiple choice questions

a.   

Computer:  
*Qu’est-ce que Patrice va faire au moment où il termine son message ?*

‘What is Patrice going to do after he finishes his message?’

Student:  
*C. Faire sa valise.* ‘Pack his bags.’

Computer:  
*C’est correct.* ‘That’s right.’

*Très bien.* ‘Good job.’

b.   

Computer:  
*Qu’est-ce que Patrice va faire au moment où il termine son message ?*

‘What is Patrice going to do after he finishes his message?’

Student:  
*A. Dîner.* ‘Eat dinner.’

Computer:  
*Non, ce n’est pas correct.* ‘Sorry, that’s incorrect.’
Il va faire sa valise. ‘He’s going to pack his bags.’

In all cases, all feedback on answers for these two question formats was entirely unenhanced.

3.7.3 Variations Among Treatment Conditions

Beyond the features outlined above, the lesson was configured into eight unique versions (see Figure 8 below), the differences being in the way the program would respond to the short answer questions (i.e., those which elicited the prepositions being targeted) and whether the target prepositions were enhanced.

<table>
<thead>
<tr>
<th>Input Enhancement</th>
<th>Corrective Feedback Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FoM</td>
</tr>
<tr>
<td><strong>- Enhanced</strong></td>
<td>Version 1 (= Condition 1)</td>
</tr>
<tr>
<td><strong>+ Enhanced</strong></td>
<td>Version 2 (= Condition 2)</td>
</tr>
</tbody>
</table>

FIGURE 8. Lesson versions by corrective feedback type and input enhancement level.

The following sections provide an illustrated explanation of how different versions of the lesson handled the various eventualities.
3.7.3.1 Feedback on Correct Short Answers

For correct student responses to the short answer questions, all eight versions of the program responded identically (i.e., Très bien. ‘Very good.’). Note, however, that the word correct here must be qualified, for although Versions 3 through 8 of the lesson required both the right location and the right preposition (e.g., aux Etats-Unis), Versions 1 and 2 would accept any answer that merely included the former (e.g., Etats-Unis, les Etats-Unis, aux Etats-Unis, en Etats-Unis, à l’Etats-Unis, etc.).

To avoid providing negative feedback because of spelling and gender errors, the learners were given explicit instructions on the use of proper writing conventions as well as a map for reference (see Figure 9). Incidentally, despite formal indications to the contrary, the program was not case-sensitive, although some answers would not fit in the field with the caps lock on. Figure 10 shows an example of the computer handling of a correct response for all eight instructional conditions.
FIGURE 9. Short answer question prior to student response.
3.7.3.2 Feedback on Factually Incorrect Answers

On the other hand, if the computer encountered a student answer with the wrong geographical location, odd-numbered versions of the lesson would simply reject the given response and provide an appropriate alternative (14a), whereas even-numbered versions would additionally highlight the place name in bold (14b).

(14) Unenhanced and enhanced computer-student interactions with incorrect geographic places
Thus, boldfacing in the enhanced conditions was the only difference in the way the various lesson versions handled feedback involving factually incorrect responses, for the correction always began with an explicit rejection and ended with an exemplary model (see Figures 11 and 12 below).
FIGURE 11. Feedback to factually incorrect response on short answer question for odd-numbered lesson versions (i.e., 1, 3, 5, and 7).
FIGURE 12. Feedback to factually incorrect response on short answer question for even-numbered lesson versions (i.e., 2, 4, 6, and 8).
3.7.3.3 Feedback on Factually Correct but Grammatically Incorrect Answers

Philp (2003) notes that individual differences in working memory capacity and current level of language development may also play a role in learners’ ability to notice and effectively utilize recasts. She found that the length of the recast and the number of changes it involved were significant factors, with recasts of five or fewer morphemes and no more than two differences from the original utterance being easier to remember. As such, this consideration was incorporated into all the various feedback conditions, each of which is illustrated below.

3.7.3.3.1 Versions 1 and 2 (Meaning-focused Feedback, +/- Input Enhancement)

As Versions 1 and 2 of the lesson only evaluate the factual content of the answer, they both consider *en Etats-Unis* in (15) to be correct.

(15) Computer-student interactions with no form-focused correction

Computer:  *Où va Patrice pour la première étape de son voyage?*
            ‘Where is Patrice headed on the first leg of his journey?’

Student:   *en Etats-Unis*
            ‘in the United States’

Computer:  *Très bien.*
            ‘Good.’

Consequently, both display only positive feedback (see Figure 13 below).
FIGURE 13. Feedback to factually correct but grammatically incorrect response on short answer question for lesson versions 1 and 2 (FoM).

A common problem with previous studies of this type, however, is that they have failed to control the target input. Consider the two exchanges in (16) and (17):

(16)  Teacher-student exchange providing no form-focused correction

Teacher:  *Pour la première étape de son voyage, Patrice va aux Etats-Unis.*
‘On the first leg of his journey, Patrice is going to the United States.’

Où va *Patrice pour la première étape de son voyage ?*
‘Where is Patrice going on the first leg of his journey?’

Student :  *En Etats-Unis.*
‘In the United States.’
Teacher: *Très bien.*

‘Good.’

(17) Teacher-student exchange providing form-focused correction

Teacher: *Pour la première étape de son voyage, Patrice va aux États-Unis.*

*Où va Patrice pour la première étape de son voyage?*

‘On the first leg of his journey, Patrice is going to the United States.’

‘Where is Patrice going on the first leg of his journey?’

Student: *En États-Unis.*

‘In the United States.’

Teacher: *C’est ça. Aux États-Unis.*

‘That’s right. To the United States.’

Notice that in both cases, the teacher models *aux États-Unis* before the question.

However, in the second example, he or she not only provides negative evidence by reformulating the student’s grammatically incorrect response but effectively exposes the student to a second model as well. In short, the student in (17) receives twice as much input as in (16); thus, correction is confounded with exposure. As the point of Versions 1 and 2 of the experimental lesson was to provide positive evidence only, they needed to control for exposure in cases of factual correctness but grammatical error, as in (17). Thus, after the initial feedback the program gave the learner an open-ended follow-up question and a choice of several response options as in (18), a screen shot for which is provided in Figure 14.
(18) Open-ended multiple-choice question

Combien de francophones connaissez-vous aux Etats-Unis?
‘How many French speakers do you know in the United States?’

Aucun. ‘None.’
Un ou deux. ‘One or two.’
Plusieurs. ‘Several.’
Beaucoup. ‘A lot.’

À Vous

Directions: Click on your personal response to the question before continuing.

Combien de francophones connaissez-vous aux Etats-Unis?

○ Aucun.
○ Un ou deux.
○ Plusieurs.
○ Beaucoup.

Continue

FIGURE 14. Follow-up providing second model of target preposition for ‘correct’ response on short answer question.
One issue left to be resolved, however, is the provision of salient input in a grammatically non-corrective manner. Remember that in Leeman’s (2003) study, the learners were exposed to models featuring unnatural stress and intonation immediately after completing their pretest task, which may have sensitized them to their earlier mistakes. As such, highlighting the target preposition in the open-ended question would have been undesirable, for it arguably could have alerted the learners to a discrepancy with the output they had just produced. The solution, then, was to enhance the target forms in all the reading passages instead (see Figure 15 below). Thus, the highlighted preposition always appeared in the input before its relevant elicitation, not afterward in the feedback.
FIGURE 15. Enhanced reading passage from lesson version 2: Postcard 1 (from France), modeling *aux Etats-Unis* in bold red and underlined.

As for why salience in this case consisted of an underscore and a color change (to red) as well as boldfacing as opposed to boldfacing alone, the latter seemed insufficient here, where the target word is much harder to pick out because of the relatively dense surrounding context.

3.7.3.3.2 Versions 3 and 4 (Implicit Form-focused Feedback, +/- Input Enhancement)

For Versions 3 and 4, the implicit corrective feedback consisted of a recast (see
Figures 16 and 17). In short, the computer affirmed the factual content of the learner’s answer and then provided a grammatically correct model, as in (19).

(19) Computer-student interactions with recasts of incorrect target prepositions

Computer: Où va Patrice pour la première étape de son voyage?
‘Where is Patrice headed on the first leg of his journey?’
Student: en Etats-Unis
‘in the United States’
Computer: Oui, aux Etats-Unis.
‘That’s right, to the United States.’

FIGURE 16. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 3 (unenhanced recast).
3.7.3.3 Versions 5 and 6 (Non-metalinguistic Explicit Form-focused Feedback, +/- Input Enhancement)

Versions 5 and 6 provided the learners with an unequivocal rejection of their answer followed by a grammatically correct model, as in (20), screen shots for which are provided in Figures 18 and 19.

(20) Computer-student interactions with explicit rejection of incorrect target prepositions
    Computer: Où va Patrice pour la première étape de son voyage?
‘Where is Patrice headed on the first leg of his journey?’

Student: en Etats-Unis

‘in the United States’

Computer: Non, ce n’est pas tout à fait correct. ‘No, that’s not quite right.’

→ Aux Etats-Unis. ‘To the United States.’

FIGURE 18. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 5 (unenhanced rejection).
FIGURE 19. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 6 (enhanced rejection).

3.7.3.3.4 Versions 7 and 8 (Metalinguistic Feedback, +/- Input Enhancement)

Versions 7 and 8, like Versions 5 and 6, included both an explicit rejection of the learner’s answer and a grammatically correct model. However, they also inserted a metalinguistic explanation in English in between (see Figures 20 and 21).
FIGURE 20. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 7 (unenhanced rejection with rule explanation).
Short Answer

**Directions:** Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

A a É é È è Î î Ì ì Ï ï Ò ô Ì ì Ï ï Ò ô

Où va Patrice pour la première étape de son voyage?

en Etats-Unis

Non, ce n’est pas tout à fait correct.
‘Les Etats-Unis’ is grammatically plural, as are most countries whose names end in ‘s’. To express the idea of going ‘to’ or being ‘in’ a grammatically plural place, you need to replace ‘les’ with ‘aux’.

--> Aux Etats-Unis.

FIGURE 21. Feedback to factually correct but grammatically incorrect response on short answer question for lesson version 8 (enhanced rejection with rule explanation).

The texts for these materials were revised and edited by several French- and other francophone-educated language teachers in the Department of French and Italian and then trialed by several others to verify their naturalness and functionality. The entire lesson and all its variations can be found in Appendices G and H.
3.8 Procedures

In the last part of the staff meeting the week before the start of the 2006 Fall semester, the Director of the Basic Languages Program in French and Italian set aside a time for a brief explanation of the research project, which had already been approved for implementation across all sections of French 102. So that they would not be unduly influenced in the conduct of their classes, the instructors were simply told that the project would involve the implementation of some newly developed computerized reading materials and a series of cloze tests to measure their effectiveness. It was also explained that since the students would probably be unfamiliar with the cloze testing format, they would be given two practice tests as well. Next, following a short description of the general cloze test procedure, the instructors received copies of both practice tests and both forms of the research instrument, which they were asked to complete on their own and return to the principal investigator’s departmental mailbox. Finally, they were each provided with a copy of the project implementation schedule, the realized version of which appears below in Figure 22.
FIGURE 22. Project implementation schedule.

Each of the 12 course sections received its first visit in the last 25 minutes of one lesson during the second week of class. At that time, after a brief self-introduction by the researcher and an overview of the project, the students were given a short explanation of the cloze procedure (see Appendix A for script). Next, they were presented with a sentence
containing five example items as part of the instructions for a 30-deletion practice test (see Appendix B), which was administered in the 15 minutes of class time that remained. All students took the same version of the practice test, the passage for which was adapted from Chapter 9 of the 1st semester workbook and lab manual accompanying their textbook (cf. Siskin, Williams-Gascon, Virtue, & Leahy, 2003).

The following week, on the second visit to each class, the student papers for the first practice test were returned and the keyed responses were disclosed one by one on the overhead projector. The students were also given acceptable alternatives based on their actual answers and then invited to ask questions. At this point, it was emphasized that the results of this type of test needed to be evaluated differently from standard achievement tests, and it was suggested that a score of 13 out of 30 possible points would be very good for a first attempt at this level. The students were then told that they would take one more practice test that day before the start of the formal project, the schedule for which was announced immediately afterward. Finally, a second 15-minute, 30-deletion practice test (see Appendix C) was administered. Once again, all students took the same version, the passage for which this time was adapted from Chapter 10 of the same workbook.

As the pretest on the linguistic target structures was double the length of the practice
tests, to maximize the students’ disposable testing time in a minimum number of classroom visits, the second practice test papers were given back to the instructors to return to their students along with an answer key and some additional relevant information (see Appendix D) for them to post on the course management program, Desire2Learn (D2L). Pretesting began the following day. For each class, enough copies of each pretest form were counted off for half the number of students in attendance. For example, if there were 20 students present, 10 copies of each of the two pretest forms were pulled. The test papers were then shuffled thoroughly before being distributed. In case of an odd numbers of students, a coin toss was used to determine which of the two forms would be the first extra. To ensure equal numbers of both forms throughout the sample, the next class with an odd number of students present would automatically get one more of the opposite form. The students had 25 minutes to complete their assigned version of the 60-item instrument. In the session prior to taking the test, the students were told that it would count as extra credit equivalent to one free quiz grade for those answering at least 23 items correctly.

One week after their pretest, the students all came to the computer lab for the instructional treatment. Each of the eight lesson versions had been installed on different profiles on a network server, and each profile had a unique username and password which
were unknown to the students until the day of the treatment. These usernames and passwords were printed off on strips of paper and bundled together in groups of eight, one for each profile, and then shuffled. When the students arrived at the lab, each received a single strip of paper corresponding to one of the eight profiles. To ensure uniform numbers across conditions for the entire sample, in the case of uneven multiples of eight students, the remaining profiles were assigned first to the following class before the next bundle of eight was opened.

Once inside the lab, the students could sit at the computer of their choice and sign on. To begin the lesson, they clicked on an icon called “French Reading Lesson” in the center of their screen and entered their personal identifying information, which was necessary to verify their attendance. Next, they read a set of general instructions in English (see Figure 23) before moving on to the lesson, which was self-paced. They then worked through the 14 reading passages one by one, answering the questions and receiving feedback on each as the computer recorded their keystrokes and clock times. Total time spent on the lesson (in minutes and seconds) ranged from 16:31 to 48:38 (mean = 31:33), not including an eight-question survey of the students’ personal characteristics and their opinions of the learning materials at the end (see Appendix I). A 4 x 2 factorial ANCOVA at an alpha level of .05
with pretest scores as the covariate and testing order as a blocking variable revealed no
significant differences in treatment time for either corrective feedback type, $F(3, 126) = .64,
p = .594$, target form enhancement, $F(1, 126) = 1.86, p = .175$, or the interaction between the
two, $F(3, 126) = .30, p = .828$. (For details on the choice of this statistical model, see Chapter
4.) Upon completing the survey, the students were free to leave the lab. Apart from
attendance, this activity was not graded for course credit.

![Introduction](image)

**Introduction**

You are about to read the story of Patrice, a French university student who is traveling the world on a
semester at sea. The story is told as a series of postcards Patrice writes to his girlfriend as he leaves each
of 14 ports. After each postcard, you will be asked up to four related questions in a combination of fill-in-
the-blank, true/false, and multiple choice formats. Some of these questions will be factual. For these, you
will need to check your answer and receive feedback on each one before you can move on. Other
questions will ask for a personal response, which will be recorded for later reporting to your instructor
when you continue to the next screen.

In any case, you will NOT be able to refer back to any of the reading passages once you have seen the
first related question, so you should read carefully as you go. Remember, however, that your ultimate goal
is to complete all 14 passages by the end of the class period, so you will also need to be mindful of your
pace. Although this activity will not count toward your grade, it is intended to give you valuable practice
and, thus, help you improve your eventual test performance. As such, we hope that you will do your best
to make the most of this opportunity.

At the end of the main lesson, there is an 8-question survey, through which to provide feedback on
today's learning experience so that we can better tailor future lessons to students with similar backgrounds
and interests. We hope that you will answer the survey thoughtfully. We thank you in advance for your
cooperation. Bon courage!

**FIGURE 23.** Student instructions for treatment lesson.
As a single class period did not allow enough time for both the treatment lesson and the immediate posttest, the latter was administered two days later, once again in the students’ regular classrooms. The reason for this 48-hour interval was to maintain consistency across all sections, one of which met only twice per week. Although treatment effects may have been attenuated by this delay, this issue was considered unimportant for the purposes of the present investigation as its primary interest was to measure not learning, but acquisition, which connotes long-term maintenance of the target forms.

For the immediate posttesting session, the test papers were pre-printed, each student receiving the opposite form to that of their pretest (e.g., those who previously took the *francophone* version now took the *musulman* version). Prior to the day of the test, the students were told that it would count as a regular quiz, but graded on a norm-referenced four-point scale (‘exceptional’, ‘above average’, ‘average’, and ‘below average’). As before, they had 25 minutes to complete the 60-deletion exercise.

Once the pre- and posttests were scored, the results were sent to the regular classroom teachers, who posted them on D2L. Actual test papers were not returned, but students interested in reviewing their answers were invited to make an appointment outside of class. None took advantage of this offer.
Ten weeks after the immediate post-treatment measure, the students took one final cloze test, in the same form as their original pretest. Once again, the students were told that it would count as a norm-referenced quiz, which they had 25 minutes to complete. Likewise, they received their results through D2L but were asked to make an appointment if they wanted to review their actual test paper, and no one made such a request. The following week, the Director of the Basic Languages Program in French and Italian visited each classroom to solicit the students’ permission to use their data for analysis, to which over 90% of those eligible consented (see Appendix J for a copy of the script).

3.9 Data Processing

As the cloze exercises had previously been scored for classroom purposes, the data from each testing session already existed in one of three separate databases in version 14.0 of Statistical Packages for the Social Sciences (SPSS). Nevertheless, after human subjects permission was acquired, some additional processing was necessary to prepare this data for the analyses of principal interest to the main project. First, any students who either did not sign the consent form or did not complete the treatment lesson and all relevant testing were purged. Next, after the non-target cloze items were dropped, the total number of correct responses to the target items was calculated for each test administration, then the three
databases were merged. Later, a data processing specialist generated a list of random ID numbers, which were then matched to the student names in the new dataset.

The data specialist also did all the initial handling of the data from the files that the students generated while completing the treatment lesson. First he assigned the same random ID numbers to the individual student lesson files before stripping them of all other personal identifying information. He then wrote a Java script program compiling these individual files into a single text file, which were imported into SPSS. Finally, the lesson data and cloze data files were merged, the unmatched cases from the former and the student names from the latter being dropped.

Ordinarily, the integrity of the data would have been checked by more efficient means (e.g., examination of descriptive statistics, comparison of unusual data entries against original test forms), but an initial inspection of the pre-merge data files revealed that at least one entire case in one file had been unmistakably overwritten with another. At first, this error was attributed to human carelessness and, as such, the relevant data were simply re-entered. However, upon saving the file, it immediately became clear that something else was amiss when these newly entered data reverted to their previous erroneous values. Later it was discovered that the version of SPSS being used had a bug and required a patch, which was
quickly downloaded and installed. Unfortunately, in the meantime, given the possibility of extensive corruption due to an undetected virus, the data files were all saved in an alternate format and each of the approximately 10,000 data entries was individually re-checked against the original paper test forms and re-keyed as necessary.

3.10 Summary

This study identified three principal questions. First, it asked whether novice L2 French learners would show more accurate production of à, au, en, and aux with geographic place names after experiencing a set of instructional materials designed to expose them to these structures. It then also asked whether any combinations of four types of corrective feedback and two levels of input enhancement would lead to differences in either short-term or long-term performance with the target forms.

Based on the findings of previous studies, it was hypothesized that these learners would indeed display at least temporarily improved accuracy with the target prepositions after the instruction. Furthermore, those in the explicit form-focused feedback groups would exhibit greater accuracy than those in implicit form-focused conditions, who in turn would outperform those receiving only meaning-focused feedback. Finally, typographically enhanced target forms would lead to better accuracy as compared to their unaltered
counterparts, though the size of this effect would depend on the associated type of corrective feedback.

In a pretest-posttest-delayed posttest experimental design, 136 university French 102 students were randomly assigned to one of eight computerized treatment lessons, each exemplifying a single combination of corrective feedback type and input enhancement level. The next chapter will examine the results of the linguistic outcome measures in light of the three research questions and provide statistical decisions with respect to the relevant hypotheses.
CHAPTER 4

RESULTS AND ANALYSES

This chapter presents an analysis of the data collected to answer the three main questions of this study, namely whether novice L2 French learners experiencing special instructional materials incorporating the prepositions à, au, en, and aux before geographic place names showed improved accuracy on these target forms over time, whether there were any immediate differences in target form accuracy among learners receiving different types of corrective feedback and levels of typographical input enhancement, and whether there were any long-term differences in target form accuracy among learners receiving different levels of these same two variables. As described in Chapter 3, the outcome measures consisted of two randomly assigned versions of a 60-deletion cloze test containing 24 target items and 36 distractors. Thus, once the latter were removed, the possible scores ranged from 0 to 24, with one point being awarded for each correct preposition. Below is a sequential examination of the results on these measures as they pertain to each research question.

4.1 Research Question #1: Learnability of the Linguistic Target Structures

As no previous study of the L2 development of the à/au/en/aux distinction was found
in the published literature, regardless of one’s position on the issues of input enhancement or corrective feedback, it was uncertain whether the learners in the present investigation would show any observable growth over the 11-week period it encompassed. Consequently, an initial analysis was performed to determine whether learner accuracy on the target structures after the instructional treatment had improved overall, that is, independently of differences in learning conditions. Table 3 below gives the means and standard deviations of the accuracy scores for each of the three cloze testing sessions.

<table>
<thead>
<tr>
<th>Time</th>
<th># of Items</th>
<th>Mean # Correct</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>24</td>
<td>4.5</td>
<td>2.83</td>
</tr>
<tr>
<td>Immediate posttest (48 hrs.)</td>
<td>24</td>
<td>7.4</td>
<td>4.23</td>
</tr>
<tr>
<td>Delayed posttest (10 wks.)</td>
<td>24</td>
<td>8.5</td>
<td>4.60</td>
</tr>
</tbody>
</table>

From these data, it appears that the learners’ accuracy not only improved immediately after the treatment, but that it continued to grow throughout the entire period of the
investigation, even without subsequent intervention. Figure 24 illustrates this tendency.

FIGURE 24. Overall mean target form accuracy scores over time.

To determine the statistical significance of these results, a single repeated-measures analysis over all three testing sessions was at first considered. However, this option was subsequently discarded as being in violation of the assumption that all measures for such an analysis are either unique or identical, for it must be reiterated here that the study participants took only two different test forms in either A-B-A or B-A-B order. As such, the decision was made to conduct two separate analyses: one of the differences between the pretest and immediate posttest results, and another of immediate and delayed posttest performance.

In both cases, a repeated-measures ANOVA was performed for corrective feedback
type and input enhancement on target item accuracy. Testing order (i.e., A-B-A vs. B-A-B) was also included in the model each time as a blocking variable so as to remove the undesirable effects of any differences in form difficulty. Moreover, the Bonferroni procedure was used to adjust the criterion for each comparison to .025 and, thus, maintain an overall family-wise alpha of .05.

As the results in Table 4 below indicate, the immediate effect of time was not only clearly significant, $F(1, 127) = 96.76, p < .001$, but large as well, $\eta^2_p = .43$. In other words, the learners indeed improved substantially in the short term.

TABLE 4. Repeated-measures ANOVA Results for Corrective Feedback Type, Input Enhancement, and Testing Order on Pretest and Immediate Posttest Target Form Accuracy

<table>
<thead>
<tr>
<th>Effect</th>
<th>Type III SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>525.228</td>
<td>1</td>
<td>525.228</td>
<td>96.755</td>
<td>&lt;.001</td>
<td>.432</td>
</tr>
<tr>
<td>Time * Feedback</td>
<td>31.519</td>
<td>3</td>
<td>10.506</td>
<td>1.935</td>
<td>.127</td>
<td>.044</td>
</tr>
<tr>
<td>Time * Enhancement</td>
<td>2.643</td>
<td>1</td>
<td>2.643</td>
<td>.487</td>
<td>.487</td>
<td>.004</td>
</tr>
<tr>
<td>Time * Feedback * Enhancement</td>
<td>8.081</td>
<td>3</td>
<td>2.694</td>
<td>.496</td>
<td>.686</td>
<td>.012</td>
</tr>
<tr>
<td>Time * Testing order</td>
<td>62.901</td>
<td>1</td>
<td>62.901</td>
<td>11.587</td>
<td>.001</td>
<td>.084</td>
</tr>
<tr>
<td>Error</td>
<td>689.412</td>
<td>127</td>
<td>5.428</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
However, it is also important to note the finding of a significant interaction for testing order, which implies different rates of change depending on which form of the outcome measure one started with. To rule out the possibility that the learners had, in fact, improved on only one and not the other of the two testing orders, an additional repeated-measures ANOVA was run for each order. Following a Bonferroni adjustment to hold the overall family-wise alpha level at .05, the results showed a significant effect for time, whichever form of the cloze test the students took first: $F(1, 58) = 62.86, p < .002$ for the French-speaking World version; $F(1, 62) = 27.59, p < .002$ for the Muslim World version. Thus, the differential improvement between learners with different testing orders does not affect the generalizability of the conclusion regarding the principal variable of interest (i.e., the main effect of time) in the original analysis.

Next, Table 5 shows that the delayed effect of time was also significant, $F(1, 127) = 12.417, p = .001$, though not as nearly as large, $\eta_p^2 = .089$. 
Thus, it appears that the learners not only maintained their initial gains over the long term but continued to improve. On this last point, however, a word of caution is in order, for the results of most typical learning experiments show decreases in performance on delayed measures. As such, one could argue that the semblance of additional growth is merely an artifact of comparing scores from the second administration of the same form as was used for the pretest.
Nonetheless, it must also be remembered that the linguistic target structures under investigation, while not naturally salient, occur frequently in normal classroom discourse; therefore, incidental exposure to them, though undoubtedly less intense thereafter, would not have stopped with the treatment. Furthermore, the implicit process of language acquisition, as distinguished from explicit language learning, is theoretically not supposed to be subject to forgetting in any case. Finally, as was mentioned in the discussion of the research procedures in Chapter 3, while the study participants were notified of their cloze test scores, they did not receive their original test papers back, and although they were all invited to review their papers privately, none took advantage of the opportunity. Thus, the likelihood of the learners later profiting from their mistakes on the pretest seems remote, especially given the fact that they were never told that they would eventually be given the same test form again.

Nevertheless, as was the case for the previous analysis, the interaction for testing order was once again significant. Thus, just as before, an additional repeated-measures ANOVA was run for each testing order, with a Bonferroni adjustment to hold the overall family-wise alpha level at .05. This time, however, while those learners who began with the Muslim World version of the cloze test showed subsequent improvement between the immediate and delayed posttest, $F(1, 62) = 24.91, p < .002$, those who began with the
French-speaking World version did not, $F(1, 58) = .19, p > .99$. As such, the generalizability of additional long-term gains is limited to those learners who first took the Muslim World test. However, as those who first took the French-speaking World test did not show a significant difference in performance between the immediate and delayed posttest, it can at least be said that they did not regress. In other words, learners of both testing orders showed evidence of long-term acquisition rather than mere short-term learning over the course of the study.

One other important point to note with regard to Table 5, though, is the significance of the interaction with feedback. In other words, not all of the experimental groups improved equally. However, these results are uninterpretable at this point as the means on the immediate pretest may have been numerically different but statistically equivalent, as analyses further below will determine.

At any rate, it has now been established that the learners overall did indeed improve on the target structures, yet the observed gains may be more attributable to some treatment conditions than to others. While this prospect will be explored shortly hereafter in the discussions of Research Questions #2 and #3, two other interesting possibilities also arise, namely whether the score differences between testing sessions was mainly due to acquisition
of a single target form (e.g., *en*) and/or collocations from the treatment lesson (e.g., *en Inde* ‘to India’, *à Pondichéry* ‘to Pondicherry’).

Tables 6 and 7 below show the item ease index (i.e., the proportion of test takers who answered each item correctly) for the French-speaking World and Muslim World cloze tests, respectively. The results for both are arranged in ascending order of difficulty on the delayed posttest. Furthermore, the cells of the 10 easiest questions for each testing session are shaded in gray.
TABLE 6. Target Item Ease Index for French-speaking World Cloze Test

<table>
<thead>
<tr>
<th>Collocation</th>
<th>Pretest</th>
<th>Immediate Posttest</th>
<th>Delayed Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>en Belgique</td>
<td>.576</td>
<td>.643</td>
<td>.727</td>
</tr>
<tr>
<td>en Afrique du Sud*</td>
<td>.546</td>
<td>.643</td>
<td>.727</td>
</tr>
<tr>
<td>en Tunisie</td>
<td>.349</td>
<td>.529</td>
<td>.621</td>
</tr>
<tr>
<td>en Chine</td>
<td>.273</td>
<td>.471</td>
<td>.576</td>
</tr>
<tr>
<td>au Cameroun*</td>
<td>.349</td>
<td>.543</td>
<td>.546</td>
</tr>
<tr>
<td>en Inde*</td>
<td>.288</td>
<td>.529</td>
<td>.546</td>
</tr>
<tr>
<td>au Canada</td>
<td>.379</td>
<td>.500</td>
<td>.455†</td>
</tr>
<tr>
<td>au Laos</td>
<td>.197</td>
<td>.400</td>
<td>.394†</td>
</tr>
<tr>
<td>à Montréal</td>
<td>.303</td>
<td>.171†</td>
<td>.318</td>
</tr>
<tr>
<td>au Sénégal</td>
<td>.182</td>
<td>.371</td>
<td>.303†</td>
</tr>
<tr>
<td>au Liban</td>
<td>.106</td>
<td>.257</td>
<td>.242†</td>
</tr>
<tr>
<td>à Miami</td>
<td>.046</td>
<td>.157</td>
<td>.242</td>
</tr>
<tr>
<td>aux îles Normandes</td>
<td>.121</td>
<td>.243</td>
<td>.227†</td>
</tr>
<tr>
<td>en Polynésie française</td>
<td>.030</td>
<td>.171</td>
<td>.227</td>
</tr>
<tr>
<td>au Viêt-Nam*</td>
<td>.091</td>
<td>.171</td>
<td>.212</td>
</tr>
<tr>
<td>à Shanghai</td>
<td>.121</td>
<td>.071†</td>
<td>.197</td>
</tr>
<tr>
<td>aux îles Marquises*</td>
<td>.030</td>
<td>.086</td>
<td>.182</td>
</tr>
<tr>
<td>aux Antilles françaises*</td>
<td>.030</td>
<td>.143</td>
<td>.167</td>
</tr>
<tr>
<td>aux Terres Australes</td>
<td>.015</td>
<td>.129</td>
<td>.167</td>
</tr>
<tr>
<td>à Pondichéry*</td>
<td>.030</td>
<td>.086</td>
<td>.167</td>
</tr>
<tr>
<td>à Sainte-Lucie</td>
<td>.091</td>
<td>.114</td>
<td>.152</td>
</tr>
<tr>
<td>aux Pays-Bas</td>
<td>.046</td>
<td>.057</td>
<td>.136</td>
</tr>
<tr>
<td>aux Seychelles</td>
<td>.015</td>
<td>.057</td>
<td>.106</td>
</tr>
<tr>
<td>à Tahiti*</td>
<td>.046</td>
<td>.057</td>
<td>.091</td>
</tr>
</tbody>
</table>

† = decrease from previous session
* = item appearing in treatment lesson
### TABLE 7. Target Item Ease Index for Muslim World Cloze Test

<table>
<thead>
<tr>
<th>Collocation</th>
<th>Pretest</th>
<th>Immediate Posttest</th>
<th>Delayed Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>en Egypte</td>
<td>.357</td>
<td>.636</td>
<td>.686</td>
</tr>
<tr>
<td>en Somalie</td>
<td>.500</td>
<td>.530</td>
<td>.671</td>
</tr>
<tr>
<td>à Los Angeles</td>
<td>.443</td>
<td>.546</td>
<td>.657</td>
</tr>
<tr>
<td>au Tchad</td>
<td>.614†</td>
<td>.591†</td>
<td>.643</td>
</tr>
<tr>
<td>en Turquie</td>
<td>.243</td>
<td>.606</td>
<td>.629</td>
</tr>
<tr>
<td>en Russie</td>
<td>.300</td>
<td>.455</td>
<td>.600</td>
</tr>
<tr>
<td>en France*</td>
<td>.357</td>
<td>.546</td>
<td>.586</td>
</tr>
<tr>
<td>en Indonésie*</td>
<td>.186</td>
<td>.409</td>
<td>.500</td>
</tr>
<tr>
<td>à Dearborn</td>
<td>.257</td>
<td>.333</td>
<td>.500</td>
</tr>
<tr>
<td>à Paris</td>
<td>.286</td>
<td>.318</td>
<td>.471</td>
</tr>
<tr>
<td>à New York*</td>
<td>.157</td>
<td>.288</td>
<td>.471</td>
</tr>
<tr>
<td>au Pakistan*</td>
<td>.371</td>
<td>.546</td>
<td>.386†</td>
</tr>
<tr>
<td>aux Etats-Unis*</td>
<td>.114</td>
<td>.182</td>
<td>.371</td>
</tr>
<tr>
<td>au Maroc*</td>
<td>.086</td>
<td>.258</td>
<td>.343</td>
</tr>
<tr>
<td>au Portugal</td>
<td>.071</td>
<td>.258</td>
<td>.314</td>
</tr>
<tr>
<td>à Londres</td>
<td>.086</td>
<td>.242</td>
<td>.286</td>
</tr>
<tr>
<td>au Japon</td>
<td>.100</td>
<td>.273</td>
<td>.257†</td>
</tr>
<tr>
<td>au Nigeria</td>
<td>.057</td>
<td>.227</td>
<td>.229</td>
</tr>
<tr>
<td>aux Emirats Arabes Unis</td>
<td>.086</td>
<td>.227</td>
<td>.143†</td>
</tr>
<tr>
<td>aux îles Canaries</td>
<td>.029</td>
<td>.121</td>
<td>.129</td>
</tr>
<tr>
<td>aux Philippines*</td>
<td>.029</td>
<td>.288</td>
<td>.114†</td>
</tr>
<tr>
<td>aux Comores</td>
<td>.014</td>
<td>.092</td>
<td>.086†</td>
</tr>
<tr>
<td>aux Maldives</td>
<td>.000</td>
<td>.106</td>
<td>.043†</td>
</tr>
<tr>
<td>à Madagascar*</td>
<td>.014</td>
<td>.061</td>
<td>.029†</td>
</tr>
</tbody>
</table>

† = decrease from previous session  
* = item appearing in treatment lesson
No statistical analyses were performed on the differences in correct response rates on individual items over time. However, a cursory inspection of the numbers shows that in almost all cases there was improvement from the pretest to the immediate posttest and again to the delayed posttest, and without exception, the proportion of correct responses was higher at the end of the term than before the treatment.

Another interesting observation that can be made is that the treatment items do not appear to have benefited differentially from the non-treatment items in terms of improved accuracy. On the French-speaking World test, for example, the same three treatment items appeared among the top 10 easiest at all three testing times. Although the picture is a bit more complicated, the same general tendency can be argued for the Muslim World version if the top 12 are taken into consideration.

It might also be tempting to suggest an acquisition order for the four variants of à. In the case of en, for example, 5 of the 6 targets were among the top 6 easiest items overall on the French-speaking World test and all 6 were among the top 8 on the Muslim World version. The distributions of à, au, and aux on the two test versions, however, are quite different. For instance, while the au items are closer to the top than the à items on the French-speaking World test, the reverse is true on the Muslim World test. In short, there are probably intricate
relationships among a number of complicating factors, including geographical knowledge, word frequency, and even phonotactic constraints. Otherwise, one might rightly wonder why the most common correct response on the Muslim World pretest was *au Tchad* (‘in Chad’), when the students are unlikely to have encountered it anywhere previously.

Eventually, the elaboration of such a developmental sequence and the contingencies that inform it will make for an interesting study, but as the purpose of this particular investigation was to examine the effects of corrective feedback and input enhancement, further analyses on this last point will need to wait. Now that significant learner accuracy gains on the target structures have been established, however, still remaining to be addressed is the issue of whether the various instructional treatment conditions produced different short-term and/or long-term effects, the discussion of which is divided into two separate parts, each corresponding to its own research question and data analysis.

### 4.2 Research Question #2: Immediate Effects of Corrective Feedback and Target Form Enhancement

This question actually comprises three parts: 1) whether some types of corrective feedback are more effective than others, 2) whether typographical enhancement of the linguistic target forms makes a difference, and 3) whether these two factors interact. For the
corresponding analyses, as it was predicted that post-treatment performance would strongly correlate with prior knowledge of the target structures, pretest scores were additionally included in the model as a covariate. Although random assignment of the participants to treatment conditions took place subsequent to and independently of the pretesting, a breakdown of the scores for each eventual treatment group appears in Table 8 below for the purposes of illustration.
TABLE 8. Raw Pretest Target Form Accuracy Scores (out of 24)

<table>
<thead>
<tr>
<th>Treatment Condition</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1: Unenhanced FoM</td>
<td>16</td>
<td>0</td>
<td>8</td>
<td>3.8</td>
<td>2.32</td>
</tr>
<tr>
<td>#2: Enhanced FoM</td>
<td>13</td>
<td>0</td>
<td>8</td>
<td>3.5</td>
<td>2.85</td>
</tr>
<tr>
<td>#3: Unenhanced recast</td>
<td>18</td>
<td>0</td>
<td>12</td>
<td>3.9</td>
<td>2.99</td>
</tr>
<tr>
<td>#4: Enhanced recast</td>
<td>17</td>
<td>0</td>
<td>12</td>
<td>5.4</td>
<td>3.00</td>
</tr>
<tr>
<td>#5: Unenhanced rejection</td>
<td>15</td>
<td>1</td>
<td>8</td>
<td>5.1</td>
<td>2.25</td>
</tr>
<tr>
<td>#6: Enhanced rejection</td>
<td>18</td>
<td>1</td>
<td>13</td>
<td>4.6</td>
<td>2.83</td>
</tr>
<tr>
<td>#7: Unenhanced rejection with rule explanation</td>
<td>21</td>
<td>0</td>
<td>13</td>
<td>4.5</td>
<td>3.20</td>
</tr>
<tr>
<td>#8: Enhanced rejection with rule explanation</td>
<td>18</td>
<td>1</td>
<td>10</td>
<td>5.2</td>
<td>2.88</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>136</td>
<td>0</td>
<td>13</td>
<td>4.5</td>
<td>2.83</td>
</tr>
</tbody>
</table>

Thus, the scores for the entire sample (n = 136) ranged from 0 to 13, with a mean of 4.5 and a standard deviation of 2.83 correct responses. Note that these scores reflect performance only on the 24 target items and not the entire 60-item instrument. A 4 X 2 ANOVA of these scores at a .05 alpha level with testing order as a blocking variable revealed no significant
differences for feedback type, $F(3, 127) = 1.33, p = .267$, or target form enhancement, $F(1, 127) = .44, p = .508$, and no interaction between the two, $F(3, 127) = 1.02, p = .386$. Thus, randomization appears to have been successful.

Scores on the immediate posttest then ranged from 0 to 22, with a mean and standard deviation of 7.4 and 4.30, respectively. Moreover, as expected, the pooled within-groups correlation for these scores and those on the pretest was strong, $r(126) = .60, p < .01$, thus justifying their use as a covariate in the analyses to follow.

Table 9 shows the immediate posttest results, adjusted for a mean pretest score of 4.5 and broken down by corrective feedback type. Even a cursory visual inspection, however, quickly reveals that the research hypotheses were not upheld as no statistical analyses are required to determine that the mean for the implicit form-focused group was higher than those of both the explicit treatment groups.
Table 9 shows a similarly unexpected result for typographical enhancement as the mean for the unenhanced conditions was higher than that for their enhanced counterparts.

Table 10 shows a similarly unexpected result for typographical enhancement as the mean for the unenhanced conditions was higher than that for their enhanced counterparts.

Table 10 shows a similarly unexpected result for typographical enhancement as the mean for the unenhanced conditions was higher than that for their enhanced counterparts.

The immediate posttest results for corrective feedback type by typographical enhancement level appear in Table 11 below.
TABLE 11. Estimated Marginal Means and Standard Errors for Immediate Posttest Target Form Accuracy by Corrective Feedback Type by Input Enhancement Level

<table>
<thead>
<tr>
<th>Feedback Type</th>
<th>Enhancement Level</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoM</td>
<td>Unenhanced</td>
<td>6.4</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>5.9</td>
<td>.92</td>
</tr>
<tr>
<td>Recast</td>
<td>Unenhanced</td>
<td>7.8</td>
<td>.78</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>8.8</td>
<td>.80</td>
</tr>
<tr>
<td>Rejection</td>
<td>Unenhanced</td>
<td>7.8</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>6.8</td>
<td>.78</td>
</tr>
<tr>
<td>Rejection + Rule</td>
<td>Unenhanced</td>
<td>7.9</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>7.1</td>
<td>.78</td>
</tr>
</tbody>
</table>

From these numbers, it is obvious that input enhancement did not produce a consistent main effect, as Figure 25 further illustrates.
However, the results to this point are merely descriptive. Table 12 gives the results of a two-way ANCOVA of the immediate posttest scores with corrective feedback type and target form salience as independent variables, testing order as a blocking variable, and pretest scores as a covariate. These figures not only confirm the preceding observations but shed light on the nature of the effects of corrective feedback in combination with input enhancement as well: At an alpha level of .05, all apparent differences between groups were arguably due to chance as there were no main effects for either feedback type, F(3, 126) = 2.20, p = .092, or input salience, F(1, 126) = .39, p = .536, and no significant interaction, F(3, 126) = .62, p = .605.
<table>
<thead>
<tr>
<th>Source</th>
<th>Type III SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sig.</th>
<th>( \eta_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrective Feedback (CF)</td>
<td>71.175</td>
<td>3</td>
<td>23.725</td>
<td>2.198</td>
<td>.092</td>
<td>.050</td>
</tr>
<tr>
<td>Input Enhancement (IE)</td>
<td>4.157</td>
<td>1</td>
<td>4.157</td>
<td>.385</td>
<td>.536</td>
<td>.003</td>
</tr>
<tr>
<td>CF * IE Interaction</td>
<td>19.969</td>
<td>3</td>
<td>6.656</td>
<td>.617</td>
<td>.605</td>
<td>.014</td>
</tr>
<tr>
<td>Pretest Score</td>
<td>755.752</td>
<td>1</td>
<td>755.752</td>
<td>70.027</td>
<td>&lt;.001</td>
<td>.357</td>
</tr>
<tr>
<td>Testing Order</td>
<td>113.795</td>
<td>1</td>
<td>113.795</td>
<td>10.544</td>
<td>&lt;.002</td>
<td>.077</td>
</tr>
<tr>
<td>Error</td>
<td>1359.823</td>
<td>126</td>
<td>10.792</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Corrected Total</strong></td>
<td><strong>2414.757</strong></td>
<td><strong>135</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Also interesting to note in Table 12, however, is the considerable effect size for feedback type (\( \eta_p^2 = .05, \) or \( f = .23 \)). In the original design, the projected number of participants was 200 (i.e., 50 per group for this particular comparison), which would have yielded a power of .85 to detect a medium-sized effect (i.e., \( f = .25 \) or greater) at a .05 alpha level (Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A., in press). Although the total French 102 enrollment was close to 200 in the semester when the study was conducted, many students unfortunately had incomplete data because of imperfect attendance over the multiple
testing sessions and thus had to be excluded. The revised power estimate with the given sample size was then .67; however, the observed effect was also smaller than that which was expected. In any case, the results of the present study do not support the hypothesis that some feedback types or levels of target form salience would be immediately more effective than others.

4.3 Research Question #3: Long-term Effects of Corrective Feedback and Target Form Enhancement

Here the same basic analytic techniques were used as for answering the preceding question. Table 13 below shows the means and standard errors for the delayed posttest scores for each group. Notice, however, that the mean score for the recast condition has now fallen from highest to lowest.
TABLE 13. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Corrective Feedback Type

<table>
<thead>
<tr>
<th>Feedback Type</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoM</td>
<td>8.6</td>
<td>.69</td>
</tr>
<tr>
<td>Recast</td>
<td>7.6</td>
<td>.62</td>
</tr>
<tr>
<td>Rejection</td>
<td>8.7</td>
<td>.64</td>
</tr>
<tr>
<td>Rejection + Rule</td>
<td>9.1</td>
<td>.59</td>
</tr>
</tbody>
</table>

It is important to remember, however, that no significant differences were found on the immediate posttest. In other words, for practical purposes at this time, the means of all four groups should be seen as statistically equivalent. Nevertheless, the analysis of the immediate posttest had also raised the possibility that the meaning-focused group had fared worse than the three others and that the present study simply lacked sufficient power to detect it.

Figure 26 illustrates the mean accuracy scores for the four corrective feedback types over all three testing sessions.
FIGURE 26. Mean target form accuracy scores by corrective feedback conditions over time.

These results are clearly much better aligned with the initial hypotheses, especially since it was predicted that unenhanced implicit feedback would not show any difference from meaning-focused feedback of any kind. In other words, implicit feedback was predicted to interact with input enhancement, and although the immediate posttest scores for the latter within feedback conditions were generally lower, this one combination was the sole exception. Table 14, however, shows that input enhancement also had a reversal of fortune, with learners in the enhanced condition this time outperforming those in the unenhanced.
TABLE 14. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Input Enhancement Level

<table>
<thead>
<tr>
<th>Enhancement Level</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unenhanced</td>
<td>8.4</td>
<td>.44</td>
</tr>
<tr>
<td>Enhanced</td>
<td>8.6</td>
<td>.46</td>
</tr>
</tbody>
</table>

So far, all the results are consistent with those that were predicted. Table 15 now provides the final breakdown of scores by enhancement level within feedback type. Once again, however, it is clear that the eight treatment groups did not fit the expected pattern.

Furthermore, while statistical significance still needs to be determined, the differences between the enhanced and unenhanced versions of the meaning-focused and explicit form-focused treatments raise a few questions. First, no significant interaction was predicted for the meaning-focused group, but the difference between the means for the enhanced and unenhanced sub-groups is larger than that for the explicit form-focused group with rule explanation. In short, if the interaction is significant for the latter, it is likely so for the former as well, although unequal numbers of participants per group and differing standard errors make this determination uncertain without proper statistical tests.
TABLE 15. Estimated Marginal Means and Standard Errors for Delayed Posttest Target Form Accuracy by Corrective Feedback Type by Input Enhancement Level

<table>
<thead>
<tr>
<th>Corrective Feedback Type</th>
<th>Enhancement Level</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoM</td>
<td>Unenhanced</td>
<td>7.8</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>9.5</td>
<td>1.03</td>
</tr>
<tr>
<td>Recast</td>
<td>Unenhanced</td>
<td>7.4</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>7.7</td>
<td>.89</td>
</tr>
<tr>
<td>Rejection</td>
<td>Unenhanced</td>
<td>9.9</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>7.4</td>
<td>.87</td>
</tr>
<tr>
<td>Rejection + Rule</td>
<td>Unenhanced</td>
<td>8.5</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>Enhanced</td>
<td>9.7</td>
<td>.87</td>
</tr>
</tbody>
</table>

Another curiosity revealed by Table 15 is the apparent order of the two means in the rejection condition, illustrated in Figure 27 below.
FIGURE 27. Delayed posttest target form accuracy by corrective feedback type by input enhancement level.

This disordinal relationship presages the results of the upcoming analysis of the main effects of input enhancement, namely their uninterpretability even should they appear statistically significant. Ironic, however, is the fact that whereas on the immediate posttest, the top scoring group was the only one for which input enhancement implied an advantage, this time the top group is the only one for which it did not.

Once again, though, to this point these results are merely descriptive and speculative. To determine the reality of any of these observed differences, another two-way ANCOVA with feedback type and input enhancement as fixed factors, testing order as a blocking variable, and pretest scores as a covariate was performed. The results appear in Table 16.
below. Once again, however, at an alpha level of .05, there was no significant effect for either feedback type, $F(3, 126) = 1.15, p = .334$, or input enhancement, $F(1, 126) = .08, p = .774$, and no interaction, $F(3, 126) = 2.15, p = .097$.

**TABLE 16. ANCOVA Results for Delayed Posttest Target Form Accuracy**

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sig.</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrective Feedback (CF)</td>
<td>46.025</td>
<td>3</td>
<td>15.342</td>
<td>1.145</td>
<td>.334</td>
<td>.027</td>
</tr>
<tr>
<td>Input Enhancement (IE)</td>
<td>1.108</td>
<td>1</td>
<td>1.108</td>
<td>.083</td>
<td>.774</td>
<td>.001</td>
</tr>
<tr>
<td>CF * IE Interaction</td>
<td>86.458</td>
<td>3</td>
<td>28.819</td>
<td>2.151</td>
<td>.097</td>
<td>.049</td>
</tr>
<tr>
<td>Pretest Score</td>
<td>876.162</td>
<td>1</td>
<td>876.162</td>
<td>65.394</td>
<td>&lt;.001</td>
<td>.342</td>
</tr>
<tr>
<td>Testing Order</td>
<td>36.796</td>
<td>1</td>
<td>36.796</td>
<td>2.746</td>
<td>.100</td>
<td>.021</td>
</tr>
<tr>
<td>Error</td>
<td>1688.175</td>
<td>126</td>
<td>13.398</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Corrected Total</strong></td>
<td><strong>2857.735</strong></td>
<td><strong>135</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Despite the nonsignificant findings, interesting to note among these results is the substantial drop in the effect size for feedback (now $\eta_p^2 = .027$, or $f = .17$) since the first posttest. Recall that the original recruitment target was 200 participants, which would have
greatly increased the probability of concluding in favor of a significant effect for feedback immediately following the treatment, even though the actual effect size was smaller than expected. However, as in many previous studies, this apparent effect would have dissipated after a few weeks. While post hoc power analyses have been criticized on other grounds (cf. Yuan & Maxwell, 2005), the point here is that, ironically, a larger number of participants might have led to a claim for a temporary effect (i.e., learning, as opposed to acquisition) that would only have become evident because of the delayed measure, which many studies do not include. In short, it is imperative to have not only adequate power but sufficient time as well.

4.4 Summary

Through the preceding series of statistical analyses, each of the three main research questions of this investigation was examined in turn. First, comparisons of the pretest, immediate posttest, and delayed posttest scores indicated that participants in the experiment overall improved their accuracy in producing the linguistic target forms, not only 48 hours after the treatment but 10 weeks later as well, thus providing evidence of the learnability of the selected structures. However, subsequent analyses of the immediate posttest results showed no significant effects for corrective feedback type or input enhancement, nor for the interaction. Thus, the research hypotheses predicting at least short-term relative benefits of
some combinations of these two variables were not confirmed. Moreover, while it initially
seemed possible that implicit form-focused feedback had an important effect that the current
experiment was underpowered to detect, further analyses of the delayed posttest revealed that
this effect was temporary at best, for again there were no significant differences among any
of the treatment groups for either of the dependent variables or the interaction between them.
The next chapter will offer an interpretation of these findings and suggest implications for
teaching as well as future research.
To correct or not to correct? This question has been a source of debate in second language instruction since the inception of the field. While nativists argue against any role for negative evidence in L1 acquisition, many SLA researchers see L2 acquisition as a separate process and, thus, discount this view.

Previous studies attempting to unequivocally demonstrate the usefulness of corrective feedback in SLA, however, have generally failed. Moreover, input enhancement has once again made its way into the discussion along with a renewed interest in the effects of frequency and salience.

This study attempted to tease apart the effects of input enhancement and corrective feedback on university-aged novice L2 French learners’ acquisition of a set of communicatively redundant morphemes (i.e., à, au, en, and aux). Following a maximally 50-minute computerized lesson instantiating one of eight combinations of the two independent variables, the study participants took alternate versions of a cloze test, one after 48 hours and the other after 10 weeks, to determine their gains on the target structures. Below is a summary of the findings and their implications for language pedagogy as well as a discussion.
of their limitations and some suggestions for future research.

5.1 Synopsis of Findings

The present study focused on three main questions: 1) whether the participating learners in general showed post-treatment improvement in accuracy on the target forms over time, 2) whether different types of corrective feedback and/or levels of typographical input enhancement led to short-term differences in accuracy on these target forms, and 3) whether different levels of these two variables produced such differences in the long term.

With respect to the first question, statistical analyses showed that the mean target form accuracy score for the entire sample was significantly higher immediately after the treatment than before. Thus, the target structures were indeed within the learners’ developmental range. In fact, further growth on these structures was also evident 10 weeks later with no additional intervention in between, though the magnitude of this second increase was substantially lower than the first, despite the much longer time interval. In any case, maintenance of the initial gains alone provides support for the claim that these changes were due to acquisition rather than mere conscious learning.

The analyses pertaining to the second question, however, yielded a few unexpected results. For example, the mean target form accuracy scores on the immediate posttest were
lower for learners in the enhanced than in the unenhanced conditions. Furthermore, the
highest obtained mean among the four feedback groups was that of the implicit form-focused
group. However, these findings and all other between-group differences were attributable to
chance, not only on the immediate posttest but on the delayed one as well.

Thus, as learners in none of the treatment conditions showed statistically significant
differences in post-treatment accuracy in either the short or long term, it appears that
comprehensible input alone was enough to spur development of the target forms. Below is a
discussion of the implications for language teachers.

5.2 **Pedagogical Implications**

The good news for those who oppose the use of corrective feedback is that even the
learners in this study who received comprehensible input alone showed statistically
significant improvement in accuracy on the target forms up to after 10 weeks after the
treatment, although the changes over the intervening period may not be striking to the casual
observer. These findings are consistent with Ayoun’s (2005) conclusions that, given enough
time, L2 learners may eventually match the performance of L1 users on certain linguistic
tasks. Unfortunately, most L2 learners probably do not have that much time. For example,
US Foreign Service officers learning French require an average of eight weeks and 240 hours
of instruction to achieve a rating of 1 or 1+ on the Foreign Service Institute Oral Proficiency Interview (Liskin-Gasparro, 1982). One can only wonder what to expect from university learners who log only approximately 100 class hours over an entire academic year. Since corrective feedback (and even input enhancement) in this study did not appear to have any effect on beginners learning the target structures under investigation, clearly neither of these provides an immediate solution to the problem of foreign language learning.

Nonetheless, even if corrective feedback should never prove effective in producing changes in linguistic behavior, it may still merit consideration as a pedagogical technique, for it also has other useful functions such as providing reinforcement for current effort levels as well as incentive for their increase (Annett, 1969). Good and Brophy (2000) add that feedback provides immediacy and impact to an activity, arguing that most students find it psychologically more difficult and less rewarding to go back and re-do an activity or task once they have completed it than to respond to immediate feedback as they proceed the first time.

As was mentioned in the description of the methodology, at the end of the treatment, the learners completed a survey on their experience. While the results were not included in the analysis chapter because of their tangential relationship to the main research questions,
here they become more relevant. Granted, as the survey was tied to the treatment data, it was not anonymous. Furthermore, since the learners rarely, if ever, used computers in their regular lessons, novelty may also have affected the responses. Nevertheless, on a five-point Likert scale (1 = much less enjoyable, 5 = much more enjoyable), learners across all conditions largely (67%) indicated that they found the computerized lesson either somewhat more or much more enjoyable than traditional paper-and-pencil exercises. A 4 x 2 ANCOVA using the same model as for the main analyses of this study showed that the average score on this item (M = 4.0, SD = 1.36) was significantly higher than the expected value of 3.0, F(1, 126) = 33.54, p < .001. Moreover, there was no significant difference in scores for either corrective feedback type, F(3, 126) = .59, p = .621, target form enhancement, F(1, 126) = 1.70, p = .194, or the interaction between the two, F(3, 126) = 1.16, p = .329. In other words, all learners appeared to have preferred the computerized lesson, regardless of experimental treatment condition.

Of course, other factors such as visual appeal may also have played a role in shaping these learners’ opinions, but Good and Brophy note that feedback is among the reasons for the popularity of recreational video games, the sexier cousins of most computer-based instruction modules. In fact, while computer-assisted language learning (CALL) has now
evolved far beyond its initial drill-and-kill phase, the computer can always reprise its former role as tireless provider of consistent feedback if necessary, and by taking over controlled practice-type activities such as that of the treatment lesson, it may even free up class time for more open-ended communication activities.

Moreover, within the L2 classroom, there are a few other perspectives on the use of corrective feedback that also merit consideration, the first of which is learner beliefs and expectations. Fotos (1998) notes the difficulties of implementing non-traditional teaching approaches in some contexts abroad, for example, because of an expected emphasis on grammar instruction and error correction. In support of this claim, Schulz (1996, 2000) collected survey data in the US and Colombia to compare L2 teacher and learner attitudes toward these two traditional pedagogical staples and found that indeed learners in both cultures had the more positive views with respect to each. Although a somewhat smaller number of learners embraced the idea of having more formal grammar instruction, the overwhelming majority indicated that they liked and expected to be corrected. In her conclusions, she thus advocated the need to explore learners’ beliefs and to match those of the teacher with student expectations.
Another plausible option, however, would be to actively shape the latter by providing a pedagogical rationale for shifting one’s emphasis away from traditional structural issues in favor of a more communicative approach. After all, it is the teacher who should be more knowledgeable about the language acquisition process, and many students may have little experience with non-traditional methodologies on which to base their opinions.

Moreover, not all learners in every case necessarily prefer more conventional teaching practices. For instance, in a survey of intermediate university L2 French learners’ attitudes toward grammar instruction and extensive reading, Dupuy (1997) found that the latter was deemed to be both more enjoyable and more beneficial. In some of the open-ended responses, she also saw a dichotomous characterization of reading as fun and interesting (especially when self-directed), and grammar as boring, ineffective, and evanescent.

These results are not necessarily incompatible with those further above, however, if the distinction made in Chapter 2 between preemptive and responsive negative evidence is maintained. In short, one could argue that it is the pre-teaching of rules for which the students have no immediate use that is the cause of their boredom. Corrective feedback, on the other hand, is individualized and may, thus, be perceived by the students as having greater value.
Even if students do not mind grammar instruction, it may not be particularly effective. It is important to remember that this preemptive negative evidence is almost always coupled with corrective feedback, thus making its unique effects difficult to discern. Furthermore, as Pienemann (1989) argues, learners cannot mentally store structures beyond their level and later retrieve them when their development catches up. While metalinguistic feedback may also target structures prematurely, at least it is responsive to actual language use. Thus, it provides teachers with an escape from the unfortunate tendency to try to teach the entire grammar of a language at the lower levels (Ariew, 1982).

Another disadvantage of preemptive negative evidence is that it can jeopardize the intended value of subsequent communicative activities. In fact, it is possible that this misordering is responsible for lack of transfer: Learners are focused on linguistic form to the point that they filter out much of the implicit stimuli they would perceive ordinarily. Unfortunately, most instructors hardly have time to teach and have learners practice everything in their curriculum once, let alone insert an additional period to allow the learners to explore the language freely before any pedagogical intervention as well. Thus, metalinguistic feedback (like that in the rejection plus rule condition) may be a viable alternative unless and until it is shown to be problematic (e.g., a comparatively inefficient use
of time or a demotivating interruption to communication).

Ellis, Basturkmen, and Loewen (2002) distinguish two main types of form-focused (FonF) instruction: planned and incidental. The difference is that in the case of planned FonF, communicative tasks are specifically designed to elicit the use of certain pre-selected target forms. Thus, incidental metalinguistic feedback may be the best compromise for maintaining an instructional focus on communication and still providing learners with the corrective information they may desire. Moreover, there is nothing to lose. As Woods (1989) points out, if the students become more accurate in their L2 production as a result of increased sensitization, so much the better and, if not, at least it does not detract from their opportunity to learn the language through meaningful communication.

Once the issue of whether or not to provide corrective feedback is resolved, however, several other questions remain, such as which linguistic features to target and for whom. Although Schulz’s studies indicated that both American and Colombian learners across all proficiency levels desire corrective feedback, it is interesting to note that the percentage of Colombian learners expressing such sentiment was considerably greater, which raises questions about cross-cultural expectations of eventual levels of attainment. Furthermore, it is possible that learners at different levels have different ideas about the kinds of errors they...
would like to have corrected. For instance, beginning students, who are more likely to be
taking compulsory courses, may want mostly *conversational* feedback as they struggle to
communicate their basic message, whereas their intermediate and advanced counterparts,
who already have a basic command of the language and have elected to continue their studies,
may be more interested in *didactic* matters as they seek to more closely approximate L1 user
norms.

At this point, however, whether a relationship between learner proficiency levels and
preferred types of corrective feedback actually exists remains a matter for speculation (and
future research). Assuming otherwise, a more important issue might be feasibility within
situational constraints. For example, as cognitive/academic language is generally acquired
much later than social language, it would be difficult for instructors to use metalinguistic
strategies in beginning level classes without resorting to the learners’ L1, a practice which
may be in conflict with either institutional policy or personal teaching philosophy.

Foreign language (as opposed to second language) environments pose some
additional problems as well. For instance, the vast majority of L2 teachers in these situations
have learned the language with scant input themselves. Thus, they may lack the linguistic
command to provide their learners with an adequate model, which may explain the
widespread resistance to communicative methodologies in general in many foreign language contexts.

Once all these other important issues regarding corrective feedback have been resolved, one final remaining question is how to handle its implementation in terms of classroom management. Whereas there has been strong advocacy in the last two decades for a move toward learner-centered instruction, at all but the most advanced of levels, the teacher must assume the central role, for it is he or she who will have to provide the relevant feedback.

Of course, one might argue that the teacher could circulate around the room and monitor communicative pairs or groups, but interjecting recasts into students’ otherwise private exchanges would, in normal situations, be an egregious breech of protocol. In other words, outside the classroom context, such third party intrusions would be perceived as hostile; thus, in order to provide feedback, the teacher not only has to interrupt the flow of any naturalistic interaction that is taking place but must invoke special privilege to do so.

While such interventions are certainly not uncommon in typical classroom settings, their general absence beyond the confines of the learning environment clearly calls their ecological validity into question. Moreover, as only a nominal participant in the students’
conversation, the teacher would also likely send the signal that his or her main concern is in fact form, not meaning. One possible solution would be for the teacher to play an active role as a conversation partner or group participant, but this option would limit his or her mobility during the lesson, which would then make the most reliable feedback available to only a small number of students and, in turn, raise fairness issues.

5.3 Limitations of the Present Study

The computerized lesson was designed to overcome a number of the constraints of typical classroom-based studies. For example, it permitted random assignment of participants to treatment conditions, a practical impossibility in most intact classrooms. It also afforded the swift collection of a relatively large data set that otherwise would have taken multiple researchers several weeks to amass by scheduling and administering a comparable number of individual appointments. Moreover, the latter arrangement would have required the recruitment of volunteers, who may not be representative of the general learner population. Finally, in each condition, the computerized treatment provided the very same type of feedback each time the same error was made. In short, the lesson design was conceived as an attempt to increase power to detect significant effects while minimizing some of the common threats to internal validity.
One problem, however, was that the computer could not evaluate learners’ oral production. While having the learners produce written language in itself reduces the generalizability of any significant findings, it also introduced an additional design constraint that may have led to an aberration of the results. In short, the options were to employ a mixed modality (i.e., the learners type responses to audio recorded questions) or a consistent text exchange (i.e., the learners type responses to written prompts), the latter of which was eventually chosen.

The problem with this choice becomes apparent in trying to explain why input enhancement seemed to negatively affect the non-metalinguistic explicit form-focused group, especially on the delayed posttest, where the apparent difference was particularly stark. Lorch (1989) suggests that one of the benefits of typographical cues is that they provide a simple mechanism for directing searches for specific information. However, these linguistic targets were not intended merely as information but as cues themselves for detecting the abstract grammatical relationships they represent. It is then possible that the learners in this condition may simply have been ‘gaming the system’, that is, scanning the text for the bold-faced word and only processing it very shallowly.
Participants in the unenhanced version of the same feedback condition, on the other hand, had to read the feedback attentively to locate their errors and, thus, might have given them more consideration. Although there is no statistical evidence that the learners in either of these two groups spent significantly more or less time on the treatment than the other (cf. Chapter 3, pp. 164-165), had the feedback been given auditorily, learners in these both groups might have processed it more similarly. Consequently, those in the unenhanced rejection condition might have noticed fewer of the target forms, which in turn might have led to a main effect for input enhancement since there would no longer be this anomalous cell mean pulling in the opposite direction.

Unfortunately, it would take a much more labor-intensive intervention to resolve this issue, namely one-on-one synchronous interaction. Moreover, results of the post-treatment survey revealed that only 40% of participants might be interested in participating in other general French research outside of class, whereas this number rose to 58% for the piloting of similar computer-based materials. Thus, short of a clever hybrid design incorporating a computerized presentation with human feedback or a substantial advance in voice recognition capabilities, this problem does not seem to have an immediate solution.
Another fault of the present study was its failure to distinguish between the development of existing structures and the learning of new ones. N. Ellis (2002b) notes that explicit learning can be fine-tuned through input. In other words, learners who already had a rule may simply have been implicitly affirming it. In this case, the rule would have been that for the use of the four target prepositions not specifically with geographic place names, but in common locative expressions. As was mentioned in the description of these linguistic structures in Chapter 3, their usage is identical for à, au, and aux. It is the re-mapping of à la and à l’ into en that is arguably the most difficult. In short, the learners might have been using à, au, and aux but learning en, although they had no doubt encountered even the latter in numerous other contexts (e.g., en hiver ‘in winter’, en classe ‘in/to class’). In fact, Dasse (2004) found similarly disappointing results in her comparison of beginning and intermediate L2 French learners in various form-focused conditions versus a control, namely no differences among groups who had all been previously taught a similarly high-frequency structure (the passé composé).

Thus, it would have been preferable to assess explicit knowledge of both relevant applications of the target structures. However, it was not readily apparent how to do so without biasing the experiment or requesting more than the already generous seven
interruptions to the normal class schedule.

Moreover, as Bley-Vroman (1983) points out, the categories of an interlanguage may not always strictly coincide with those of the target language. For example, the learners may have come up with the understanding (whether implicit or explicit) that *en* precedes countries that normally start with *la* (e.g., *la France* → *en France*) but still erroneously applied *à l’* to other feminine countries (e.g., *à l’Algérie*). However, as the cloze procedure only permits one word per blank, it does not accurately reveal learners’ development in this case, for their natural choice of response is not allowed. While an alternate instrument may have eliminated this problem, an additional concern was how to reliably elicit substantial numbers of the target forms in a single integrative task.

As one of the cooperating teachers also noted, control of this structure requires not only grammatical but geopolitical knowledge as well. For instance, *Madagascar* is a grammatically masculine country, but it is also a small island, which is why it takes *à* rather than *au*. Interestingly, as was mentioned in Chapter 3 in the discussion of the cloze test results for the native and near-native speakers, not one of the latter, including two Africans who surely know that Madagascar is an island, put *à*. In short, cultural knowledge (i.e., the French perception of the relative size of Madagascar) may also be necessary.
While preexisting knowledge of the target structures may have reduced the magnitude of the learning effects, it is also possible that the à/au/en/aux distinction in particular can be learned implicitly, for two of the primary determinants of implicit learnability are adjacency and repetition (N. Ellis, 2002b). As Sharwood Smith (1991) notes, even simple relative frequency can make certain linguistic elements perceptually salient, although research results on the effectiveness of input floods are inconsistent (cf. Trahey & White, 1993; Williams & Evans, 1998).

Another weakness of the present study is that it did not assess to what degree participants in the metalinguistic feedback group actually learned the rules. In other words, maybe there was no difference in the types of awareness (i.e., implicit vs. explicit) learners in the various treatment conditions were able to bring to bear on the processing of the input. Unfortunately, once again, to investigate this possibility would have required an additional measure as well as instructional time in which to administer it.

Perhaps the greatest shortcoming of this study, however, was its underlying assumption that a one-shot treatment of such limited duration could produce significant results. While the learners in all treatment conditions showed sustained improvement in the accuracy of their use of the target structures, average attainment still fell far short of L1 user
norms. Furthermore, just as learner performance did not seem to benefit from either corrective feedback or input enhancement, nor did it appear to suffer. Thus, it remains possible that differences among the treatment conditions examined here might have materialized over a longer period of time. Unfortunately, the instructional treatment employed in this study arguably reflects the practical reality of most introductory level language courses: An overfull curriculum dictates that each learning objective receives a certain allotted amount of class time with little margin for review, let alone re-teaching.

5.4 Suggestions for Future Research

Before the effectiveness of corrective feedback and/or input enhancement can ever be demonstrated, first, appropriate linguistic targets will have to be identified. The choice of experimental target forms is generally based on problematicity as determined by a combination of five factors: 1) remedial, 2) grammatical complexity, 3) acquisitional sequences, 4) linguistic theory, or 5) psycholinguistic theory (R. Ellis, 2006). While the present study drew on both Factors 1 and 5, an additional suggestion would be to consider corpus linguistic data as well. For instance, it would be interesting to know the relative frequencies of not only locative expressions including à, au, en, and aux, but of other competing expressions as well (e.g., à présent ‘now’, au lait ‘with milk’, en voiture ‘by car’,
jouer aux cartes ‘to play cards’), for the associations of these forms do not limit themselves to the scope and sequence of the coursebook. Moreover, learner corpora need to be compiled and compared with L1 data to determine patterns of difficulty and their possible sources.

Once a linguistic target has been identified, the next question that needs to be answered is how to structure the experimental task so that the learner might possibly map the form-function relationship in the course of live communication. For instance, if the meaning of an utterance is unknown but contingent on its form, one of two possibilities exists. First, the learners may already know the form, in which case use occurs, not learning. Second, they do not know the form, in which case they must decipher the meaning of the utterance, realize that the form provided a cue, and correctly remember what the form was. Necessary redundancies like à, au, en, and aux thus make attractive target structures for this type of linguistic research as they do not cause communicative breakdowns triggering global comprehension strategies and drawing attention away from form. An interesting question is whether certain types of correction can be effectively delivered so as to promote deep cognitive processing (e.g., hypothesis testing) even when meaning is not critical.

However, if the meaning is already clear, why pay attention to form? In an unpublished thesis, Plann (1975) found that the L2 learners she observed had created their
own classroom dialect, which deviated from that of the teacher and which proved resistant to instruction. While it is possible that the learners have simply acquired their own nontarget-like output in such a case where the latter undoubtedly constitutes the overwhelming majority of the input they process (Terrell, 1991), power and identity may also play important roles in determining learner attention and effort, particularly when the learners share an L1 different from the teacher’s. The point here is not to try to discredit didactic feedback altogether but merely to suggest that some learners may be more or less amenable to it, depending on their desire to join an established community of target language users.

With regard to classroom foreign language learning, Vigil and Oller (1976) once asserted that “as long as the affective messages conveyed to the student are predominantly positive, frequent instances of negative cognitive feedback are not likely to do any harm, and in fact are probably essential to a high level of attainment” (p. 294). If such were true, however, a clear advantage for recasts should be evident, but it is not. Thus, a new question arises, namely how explicit feedback interacts with individual learner differences, for it is possible that some learners take it less personally than others. It also seems reasonable that motivational variables might have an effect on L2 learners’ dispositions toward didactic corrections, and it would be worthwhile to investigate whether factors such as L2 learner
self-confidence, attitude toward error correction, and general linguistic interest have an influence on ability to notice *implicit* feedback.

DeKeyser’s (1993) look at the relationship between corrective feedback and a number of L2 learner difference variables found that learners with lower anxiety and those with higher initial ability benefited more from grammatical correction. However, he also found the same trend for learners with low levels of extrinsic motivation. In addition to the need to try to replicate these results in other contexts—DeKeyser’s participants were Dutch high school students of French—a close examination of different motivation types might also be warranted. Furthermore, although corrective feedback may lead to better performance for some learners, an exploration of its effects on attitudes and motivation is essential as well, for improvement is of little value if it comes at the expense of desire to learn.

Finally, as a largely teacher-centered endeavor, corrective feedback also compels a look at what is happening with the rest of the students while the teacher is focusing on form with a few individuals, especially in large classes. For example, Mackey (1999) found that learners who were allowed to interact personally and receive their own feedback outperformed those who merely observed. Muranoi (2000), however, found that observers improved just as much as actual participants. At issue here, then, is the question of whether it
is necessary to develop classroom structures that will allow for the majority of the students to be involved in purely communicative activities while a handful are interacting with the teacher.

5.5 Final Word to Fellow Teachers

In summary, the present study showed no significant effect for either corrective feedback or target form enhancement on beginning L2 French learners’ accuracy in their production of *à*, *au*, *en*, and *aux* with geographical place names. Furthermore, although the rate of gain declined substantially after the immediate posttest, the learners continued to demonstrate growth on the target structures without subsequent formal intervention. As such, the need for form-focused instruction in some instances may be exaggerated.

Nevertheless, it is important not to overgeneralize these findings and simply conclude that instruction does not matter and that we may do as we please, for it must be remembered that the treatment in this study was of limited duration and dealt with novice language learners acquiring a small number of specific grammar structures. Furthermore, regardless of one’s personal views on its efficacy with respect to language acquisition, corrective feedback may still have motivational value, especially since many learners both want and expect it. Thus, it should not just be thoughtlessly abandoned.
On the other hand, we must also not lose sight of the fact that the primary objective of learning languages should be for their use in situations of real communication, where meaning takes precedence over form. Furthermore, our energy may be wisely spent devising ways to provide learners with maximum amounts of comprehensible input and motivating them to seek it out on their own, especially given the limited number of hours most of them ever spend inside the classroom. As it remains possible that corrective feedback benefits some learners more than others, further research needs to be conducted in this area. Until better information is available, however, perhaps the best pedagogical course to pursue would be the development of high personal standards of reflective teaching, namely through a rigorous cycle of informed planning, decisive action, attuned observation, and critical interpretation within our own particular settings, in order to reduce our risk espousing beliefs and embracing practices that are not only unscientific but utterly impressionistic.
APPENDIX A: Project Introduction Script

“My name is Paul Lyddon, and I work for the College of Humanities Instructional Computing. Some of you may know me from the language placement testing this past summer. I also teach in the Department of French and Italian. I'm here today to tell you about a special project we've got going on in the department this semester. Over the past year I've been working with Dr. Dupuy, Director of the Basic Languages Program, to develop interactive reading materials on the computer, which we will be piloting this year in all FREN 102 classes. As such, I'll be visiting your class several times over the next few weeks and then again at the end of the semester. You'll also spend one session with me in the computer lab in ML 511. One of our goals is to see whether these new reading materials are effective, and for this we use a special kind of test, called a cloze.

“Because most of you have probably never taken a cloze test before, I'm going to give you two practice tests--one today and one next week--so that your results on later ones that count will reflect only your reading ability and not your familiarity with the testing format. A cloze test resembles a fill-in-the-blank exercise, but with a few important differences. First of all, it is designed to measure reading skill, not subject knowledge. As such, it is based on an integrated text with more-or-less regularly spaced gaps that you must restore based on limited cues and your language ability, not individual sentences requiring the provision of a single fact. Next, the answers to a cloze test can be any part of speech: not just nouns or verbs, but adjectives, adverbs, pronouns, prepositions, and conjunctions as well. In fact, these words can also be contracted (e.g., c', d', l', n'). The basic rule is that only one word can go in a blank, and it must be grammatical and make sense. As such, it's possible to come up with an acceptable answer that is different from the word that was in the original text. Finally, even native speakers generally do not score 100% on a cloze test, so don't worry if you can't think of an answer for some items. The main purpose of the cloze test is to indicate whether you would be capable of reading the original text by yourself, with the help of the instructor, or not at all.

“Now I want you to try a sample cloze test. To get the best results, skim the entire passage first, then go back over it and try to fill in each blank. If you get stuck, move on; a few items will not make any difference in your score. Finally, go over the passage one last time to make sure all your answers are grammatical and make sense, and remember: only one word per blank. Good luck!”
APPENDIX B: Practice Cloze #1

**Instructions:** The following passage contains 30 blanks, where words from the original text have been deleted. Read the whole passage at least once. Then fill in each blank with a single word that fits grammatically and makes sense.

**Example:** Dans le _______ de monde, vous avez trois sortes _______ d’amis: vos amis qui vous _______ aiment, vos amis qui ne soucient pas _______ de vous, et vos amis qui _______ vous _______ haissent. –Chamfort (1741-1794)

---

**Hélène Thibaut, journaliste québécoise**

A Montréal, les amis d’Hélène Thibaut savent que c’est une femme indépendante. Elle a de l’ambition, mais _______ place toujours l’amitié avant le _______. Elle prend des décisions intelligentes pour _______ carrière, mais sa liberté est très, _______ importante pour elle.

Son père est _______ et parle français, et sa mère _______ de Toronto et est d’origine _______. Hélène est parfaitement bilingue. Elle adore Montréal _______ où on parle français, mais elle aime _______ rendre visite à la famille de _______ mère et parler anglais avec eux. _______ est dans la famille de sa _______ qu’elle a appris l’importance _______ la communication interculturelle. A cette époque, _______ pensait devenir professeur.
et faisait études en communication interculturelle. Mais après
(15) années d’études, elle a changé opinion. Elle a décidé
(16) que le lui donnerait l’occasion de voyager de mettre en
(18) pratique cette communication importante. Et les voyages, pour elle,
(20) est la liberté.
(21)
Elle a commencé travail de journaliste à Radio Canada,
(22) elle voyage pour des reportages sur politique, la culture
(23) et les sports. supérieurs ont vite compris qu’elle beaucoup de talent pour les reportages l’étranger, et elle a commencé
(25) faire des documentaires filmés sur différents du monde.
(27) (28) (29)
C’est fantastique ! Elle partout. Elle adore montrer le monde aux
(30) Québécois et parler un peu de son cher pays au reste du monde.
Answer Key to Practice Cloze #1

The first answer for each item is the word from the original text. Others listed were considered acceptable alternatives among actual student responses.

1) Elle, Hélène
2) travail
3) sa
4) très
5) québécois, français, parisien
6) vient, est
7) anglophone, canadienne
8) où
9) beaucoup, souvent, aussi, plus
10) sa
11) C’
12) mère
13) de
14) elle, Hélène
15) des, ses
16) quelques, des, ses, ces, deux, trois, etc.
17) d’, son
18) journalisme
19) et
20) si, très
21) c’
22) son, le
23) où, et
24) la
25) Ses
26) a, avait
27) à, sur
28) à, de
29) pays, personnages
30) voyage, va, travaille
APPENDIX C : Practice Cloze #2

Instructions: The following passage contains 30 blanks, where words from the original text have been deleted. Read the whole passage at least once. Then fill in each blank with a single word that fits grammatically and makes sense.

Example: Dans le (a) monde, vous avez trois sortes (b) d' amis: vos amis qui vous (c) aiment, vos amis qui ne soucient pas (d) de vous, et vos amis qui (e) vous haïssent. –Chamfort (1741-1794)

Une relation compliquée

Bruno a beaucoup d’amis et, en plus, il a trois sœurs ; alors il pense comprendre les femmes. Mais Camille, pour lui est un grand mystère. Une chose est certaine, Bruno (1) aime. Et pas comme une petite (2), pas comme une amie. Il veut (3) marier avec elle, même si, parfois, (4) ne la comprend pas. Et Camille ? (5) pense-t-elle de Bruno ? Elle (6) son charme, son humour et son (7), mais elle ne veut pas se (8) avec lui. Même si la patronne (9) restaurant pense qu’ils ont été (10), ce n’est pas très clair. Camille (11) dit « oui » à Bruno un jour, (12), pendant un voyage très romantique à Venise. (13) après, elle a eu des regrets. (14) a écouté ses idées ; il voulait (15)
petite famille, une voiture, un grand ______________ ; et elle a eu peur. Alors,  
______________ de mariage. C’est une question délicate ______________ la relation qui  
existe entre Camille ______________ Bruno.

     Mais au travail il n’ ______________ a pas de difficultés. A Canal 7,  
______________ forment le couple parfait. Ils ______________ tous les jours sur  
l’émission « Bonjour ! ». ______________ sont ensemble du matin au ______________, et  
ils s’entendent et s’ ______________ bien. Bruno est un collègue super, ______________  
s’il panique quelquefois quand Camille ______________ trop de temps à se maquiller  
______________ l’émission. Et Camille, elle aime ______________ avec un homme si  
dynamique et ______________.

     Et demain ? Vont-ils continuer ensemble, amis et collègues ? Vont-ils se séparer ?  
Comment répondre à ces questions ? Pour le moment, c’est impossible.
APPENDIX D: Desire2Learn (D2L) Follow-up to Practice Cloze #2

Dear French 102 Students,

Your instructor should return your papers for the second practice cloze test in class on Monday, 9/11. Please use the attached answer key to mark your own paper. The first answer for each item is the word from the original text. The others are ones that we found acceptable among the responses we collected from you.

As I mentioned during my last visit, this week you will have a chance to earn extra credit toward your final grade by taking one final practice cloze test. Please ask your instructor which day I will be coming by again to your section.

This time the test will be a bit longer (60 items, to be exact). As we want to be sure that as many of you as possible are able to finish, if you have any questions about past or future tests, I ask that you please email them to me so that you have the maximum amount of time in class (i.e., 25 minutes) to devote to the present test. The goal this time is a score of 23. To see how close you came on the last test, you should multiply that score by 2. Unlike the past two times, we cannot return your paper this time as we hope to use this test with other students in the future. Your instructor will post your results on D2L sometime after the scoring is complete. If you would like to see your actual paper and the correct answers, please make an appointment to visit me in my office and I will be glad to go over the test with you.

Incidentally, the piloting of the new interactive reading materials is set to begin next week. On your scheduled day, you should report directly to the computer lab in Modern Languages 511 at the beginning of class. The lesson is expected to take the entire period, so please be sure to come on time. Also, you will need your student identification number to verify your attendance, so please write it down and bring it with you if you do not have it memorized. You will not need any books, papers, or writing implements unless your instructor specifically asks you to turn something in that day.

Your first graded cloze test will take place within two days of the lab session. The format will be similar to that of the test you will be taking for extra credit this week. However, the grading will be done on a 4-point scale, based on your expected performance at the 102 level.
A second graded cloze test will then be administered toward the end of the semester to measure your progress.

In accordance with official French & Italian Department policy, make-ups will only be granted under the circumstances outlined in your syllabus. Anyone who does not take one or both of the graded cloze tests will automatically receive a score of zero on it/them.

I thank you for your cooperation and wish you the best of luck with your studies.

Sincerely,

Paul Lyddon

palyddon@email.arizona.edu
**Answer Key to Practice Cloze #2**

1) l’
2) sœur, amie, fille, copine, fleur
3) se
4) il
5) Que
6) adore, aime
7) énergie, attitude, charisme, chat, chien, intelligence, nez, physique, travail, visage
8) marier
9) du
10) fiancé, ensemble
11) a
12) autrefois
13) Mais, Et
14) Elle, Camille
15) une
16) appartement, chien, jardin, mariage, piano
17) pas
18) dans
19) et
20) y
21) ils
22) travaillent, sont, parlent
23) Ils
24) soir, travail
25) amusent, aiment, écoutent
26) même
27) met, passe, prend
28) avant, pour
29) travailler, être, manger, parler, voyager
30) intéressant, amusant, beau, bien, bizarre, bon, charismatique, charmant, chouette, comique, content, drôle, élégant, énergétique, excellent, fantastique, formidable, gentil, heureux, indépendant, intelligent, joyeux, magnifique, masculine, mystérieux, romantique, super, sympa, sympathique
APPENDIX E: Pretest/Posttest/Delayed Posttest (French-speaking World Version)

Instructions: The following passage contains 60 blanks, where words from the original text have been deleted. Read the whole passage at least once. Then fill in each blank with a single word that fits grammatically and makes sense.

**Le Monde francophone**

On estime aujourd’hui à un peu plus de 169 millions le nombre de francophones dans le monde, soit 3,2 % de la population mondiale, ce qui fait du français l’une des 10 langues les plus utilisées. En outre, le français figure parmi les quelques langues parlées aux quatre coins de la planète.

En Europe, la France mise à _______ (1), les francophones se trouvent essentiellement _______ (2) Belgique, en Suisse et au Luxembourg. _______ (3) tant que langue officielle, le _______ (4) se parle aussi à Monaco, _______ (5) îles Anglo-Normandes et au Val d’Aoste (au nord-ouest _______ (6) l’Italie). En plus, il y _______ (7) des populations francophones importantes en Allemagne, _______ (8) Pays-Bas et en Roumanie.

Le plus grand _______ (9) de francophones en Amérique du nord vit _______ (10) Canada, surtout au Québec et notamment _______ (11) Montréal, la deuxième ville francophone du _______ (12) après Paris. Quant au reste du _______ (13) américain, les communautés francophones principales se _______ (14) en Nouvelle-Angleterre, en Louisiane et _______ (15) Miami ainsi

La langue française maintient sa présence plus large en Afrique, où elle a statut officiel dans pas moins de 21 répandus à travers le continent, y compris Sénégal, en Côte d’Ivoire, au Niger, Cameroun et en République démocratique du Congo. français jouit aussi d’un certain prestige social Seychelles ainsi que dans d’autres de l’océan Indien. Cependant, c’en Algérie, au Maroc et Tunisie où le français n’est légalement privilégié, que l’on recense le grand nombre de francophones. Depuis quelques, le nombre de francophones augmente aussi Afrique du Sud suite à une importante en provenance des pays limitrophes nord.

Bien que sans statut officiel aucun pays d’Asie (y compris Moyen-Orient), le français s’utilise toujours dans états de ce grand continent. Par exemple, Liban tous les présidents ont été jusqu’à éduqués en langue française. On constate
un prestige similaire parmi les élites Laos et au
Kampuchéa ainsi qu’Viêt-Nam. Du reste, il existe encore des vestiges
ça et là comme les panneaux en français Pondichéry,
ancienne colonie française Inde, ou les sites historiques
l’ancienne concession française Shanghai, « le Paris
de l’orient », Chine.

En Océanie, le français est officiel Polynésie française, pays
d’outre-mer toujours relié la France. Donc, c’est la de l’enseignement et de l’administration Tahiti, l’île la plus grande et
plus peuplée de la collectivité ainsi qu’îles Marquises, lieu du tournage de la 4e de l’émission américaine « Survivor ». Elle est très présente à Vanuatu, archipel de la Mélanésie.

, le français s’étend même jusqu’Terres Australes d’Antarctique, zone économique exclusive française en
ressources minérales et lieu de tournage du « La Marche de l’empereur ». De ce fait, c’est vraiment une des seules langues parlées sur tous les continents.
List of Deleted Words (French-speaking World Passage)

1. part
2. en
3. En
4. français
5. aux
6. de
7. a
8. aux
9. nombre
10. au
11. à
12. monde
13. continent
14. trouvent
15. à
16. aux
17. reste
18. à
19. la
20. un
21. pays
22. au
23. au
24. Le
25. aux
26. îles
27. est
28. en
29. pas
30. plus
31. années
32. en
33. immigration
34. au
35. dans
36. le
37. plusieurs
38. au
39. present
40. aussi
41. au
42. au
43. linguistiques
44. à
45. en
46. dans
47. à
48. en
49. en
50. à
51. langue
52. à
53. la
54. aux
55. saison
56. aussi
57. Enfin
58. aux
59. riche
60. film
APPENDIX F: Pretest/Posttest/Delayed Posttest (Muslim World Version)

Instructions: The following passage contains 60 blanks, where words from the original text have been deleted. Read the whole passage at least once. Then fill in each blank with a single word that fits grammatically and makes sense.

Le Monde musulman

Ses 1,5 milliard d’adhérents font de l’islam la deuxième religion organisée du monde après le christianisme. La majorité des Arabes sont musulmans, mais ceux-ci ne constituent que 20% de la population musulmane totale. En fait, les musulmans représentent des ethnies très diverses et habitent tous les coins de la planète.

De loin, la plus grande concentration de ___________  (420 millions) est en Asie du Sud, essentiellement ___________ Pakistan, au Bangladesh et dans les états ___________ nord et de l’est de l’Inde. ___________ islam est également la religion majoritaire ___________ Maldives.


Les ___________ arabes constituent la troisième plus ___________


Ensuite, l’Europe son ensemble comprend 84 millions musulmans, dont 80% habitent Turquie et 8% dans les Balkans. une présence musulmane précolombienne îles Canaries et plusieurs siècles occupation mauresque en Espagne et Portugal, les musulmans sont peu nombreux ces pays aujourd’hui. En revanche, l’Europe en recense toujours 10 millions, répartis essentiellement France (environ 4,5 millions de musulmans sont concentrés Paris), en Allemagne et au Royaume-Uni 600 000 d’entre eux sont installés Londres.

Enfin, les estimations de la population des deux continents d’Amérique oscillent entre 3 11 millions, dont au moins 1,6 million Etats-Unis où 600 000 d’entre eux résident New York proprement dit. Par ailleurs, plus forte concentration d’Arabes du Moyen-Orient ne se trouvent ni Los Angeles, ni à Chicago, mais Dearborn, Michigan, où ils constituent 30% la population. De toute évidence, l’islam est vraiment une religion mondiale.
List of Deleted Words (Muslim World Passage)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>musulmans</td>
<td>21</td>
<td>aux</td>
</tr>
<tr>
<td>2</td>
<td>au</td>
<td>22</td>
<td>il</td>
</tr>
<tr>
<td>3</td>
<td>du</td>
<td>23</td>
<td>aux</td>
</tr>
<tr>
<td>4</td>
<td>L’</td>
<td>24</td>
<td>du</td>
</tr>
<tr>
<td>5</td>
<td>aux</td>
<td>25</td>
<td>en</td>
</tr>
<tr>
<td>6</td>
<td>compte</td>
<td>26</td>
<td>Les</td>
</tr>
<tr>
<td>7</td>
<td>en</td>
<td>27</td>
<td>ne</td>
</tr>
<tr>
<td>8</td>
<td>et</td>
<td>28</td>
<td>en</td>
</tr>
<tr>
<td>9</td>
<td>l’</td>
<td>29</td>
<td>au</td>
</tr>
<tr>
<td>10</td>
<td>aux</td>
<td>30</td>
<td>se</td>
</tr>
<tr>
<td>11</td>
<td>aussi</td>
<td>31</td>
<td>au</td>
</tr>
<tr>
<td>12</td>
<td>Du</td>
<td>32</td>
<td>ainsi</td>
</tr>
<tr>
<td>13</td>
<td>au</td>
<td>33</td>
<td>à</td>
</tr>
<tr>
<td>14</td>
<td>pays</td>
<td>34</td>
<td>musulmans</td>
</tr>
<tr>
<td>15</td>
<td>grande</td>
<td>35</td>
<td>comparaison</td>
</tr>
<tr>
<td>16</td>
<td>si</td>
<td>36</td>
<td>dont</td>
</tr>
<tr>
<td>17</td>
<td>en</td>
<td>37</td>
<td>en</td>
</tr>
<tr>
<td>18</td>
<td>au</td>
<td>38</td>
<td>de</td>
</tr>
<tr>
<td>19</td>
<td>de</td>
<td>39</td>
<td>dans</td>
</tr>
<tr>
<td>20</td>
<td>dans</td>
<td>40</td>
<td>de</td>
</tr>
<tr>
<td>21</td>
<td>aux</td>
<td>41</td>
<td>en</td>
</tr>
<tr>
<td>22</td>
<td>il</td>
<td>42</td>
<td>Malgré</td>
</tr>
<tr>
<td>23</td>
<td>aux</td>
<td>43</td>
<td>aux</td>
</tr>
<tr>
<td>24</td>
<td>du</td>
<td>44</td>
<td>d’</td>
</tr>
<tr>
<td>25</td>
<td>en</td>
<td>45</td>
<td>au</td>
</tr>
<tr>
<td>26</td>
<td>Les</td>
<td>46</td>
<td>dans</td>
</tr>
<tr>
<td>27</td>
<td>ne</td>
<td>47</td>
<td>occidentale</td>
</tr>
<tr>
<td>28</td>
<td>en</td>
<td>48</td>
<td>en</td>
</tr>
<tr>
<td>29</td>
<td>au</td>
<td>49</td>
<td>à</td>
</tr>
<tr>
<td>30</td>
<td>se</td>
<td>50</td>
<td>où</td>
</tr>
<tr>
<td>31</td>
<td>au</td>
<td>51</td>
<td>à</td>
</tr>
<tr>
<td>32</td>
<td>ainsi</td>
<td>52</td>
<td>musulmane</td>
</tr>
<tr>
<td>33</td>
<td>à</td>
<td>53</td>
<td>et</td>
</tr>
<tr>
<td>34</td>
<td>musulmans</td>
<td>54</td>
<td>aux</td>
</tr>
<tr>
<td>35</td>
<td>comparaison</td>
<td>55</td>
<td>à</td>
</tr>
<tr>
<td>36</td>
<td>dont</td>
<td>56</td>
<td>la</td>
</tr>
<tr>
<td>37</td>
<td>en</td>
<td>57</td>
<td>dehors</td>
</tr>
<tr>
<td>38</td>
<td>de</td>
<td>58</td>
<td>à</td>
</tr>
<tr>
<td>39</td>
<td>dans</td>
<td>59</td>
<td>à</td>
</tr>
<tr>
<td>40</td>
<td>de</td>
<td>60</td>
<td>de</td>
</tr>
</tbody>
</table>
Chérie,

Ça ne fait même pas deux heures depuis qu’on s’est parlé au téléphone mais tu me manques déjà. Je souhaiterais que tu puisses m’accompagner dans ce voyage et vivre cette aventure avec moi. Je promets de t’envoyer des nouvelles de chaque pays. J’arrive aux États-Unis mercredi. Je t’écrirai encore vers la fin de mon séjour là-bas. Maintenant je dois finir de faire mes bagages donc je te laisse pour l’instant. Je t’embrasse très fort.

Patrice
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Où va Patrice pour la première étape de son voyage?

Keyed response: (aux) Etats-Unis
À Vous

**Directions:** Click on your personal response to the question before continuing.

**Combien de francophones connaissez-vous aux Etats-Unis?**

- Aucun.
- Un ou deux.
- Plusieurs.
- Beaucoup.

**N.B.:** Only learners who did not receive corrective feedback on the preceding question saw this one.
Vrai ou Faux?

**Directions:** Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Patrice promet à sa copine de lui écrire tous les jours pendant son voyage.

- [ ] Vrai
- [ ] Faux

**Keyed response:** Faux
Multiple Choice

**Directions:** Click on the letter of the answer choice that best fits the question.

Qu’est-ce que Patrice va faire au moment où il termine son message?

- Dîner.
- Dormir.
- Faire sa valise.

**Keyed response:** C
Chérie,

Le voyage transatlantique était plutôt agréable. On est arrivé à New York mercredi matin. Ce jour-là, après un tour du siège de l’ONU, on a visité les quartiers de Manhattan, y compris Chinatown, Little Italy et Spanish Harlem. Les États-Unis, c’est un pays vraiment diversifié! Jeudi on a visité Ellis Island et la statue de la Liberté. Hier on a visité le musée Guggenheim dans la journée et on est allé voir «Le Roi Lion» à Broadway le soir. Notre séjour ici se termine, on part bientôt aux Petites Antilles, d’où je t’écrirai de nouveau.

Tendrement,
Patrice
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

Où est Patrice aux États-Unis?

Keyed response: (à) New York
**À Vous**

*Directions: Click on your personal response to the question before continuing.*

Est-ce que vous êtes déjà allé(e) à New York?

- Non, mais j’aimerais un jour y aller.
- Non, et je n’ai pas vraiment envie d’y aller.
- Oui, et je m’y suis bien amusé(e).
- Oui, mais ce n’était rien de spécial.

**N.B.:** Only learners who did not receive corrective feedback on the preceding question saw this one.
Vrai ou Faux?

Directions: Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Pendant son séjour à New York, Patrice est allé voir une comédie musicale à Broadway.

- Vrai
- Faux

Check Answer

Keyed response: Vrai
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

 Après son séjour à New York, où va Patrice ?

Keyed response: (aux) Petites Antilles
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller aux Petites Antilles?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Alors qu’il ne faisait que 1° à New York, ici il fait presque 25°. C’est formidable! La Guadeloupe est vraiment belle. Mardi on a visité le Fort Louis Delgrès et la Cathédrale de Basse-Terre, deux beaux exemples du patrimoine architectural français datant respectivement du 17ème et 18ème siècle. Hier, on a fait une randonnée au sommet du volcan de la Soufrière, le point culminant de l’archipel. Ce matin, on a visité le Jardin Botanique de Deshaies, un espace merveilleux appartenant autrefois au comédien Coluche. Ce soir on commence un long voyage aux îles Marquises en passant par le canal de Panama. Je t’écrirai davantage de là-bas.

Tendrement,
Patrice
Multiple Choice

**Directions:** Click on the letter of the answer choice that best fits the question.

Quel temps fait-il en Guadeloupe?

- ☐️ A  Il fait beau.
- ☐️ B  Il fait frais.
- ☐️ C  Il pleut.

Keyed response: A
Vrai ou Faux?

Directions: Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

La Soufrière est un volcan.

○ Vrai
○ Faux

Keyed response: Vrai
Où va Patrice après son passage par le canal de Panama?

Keyed response: (aux) îles Marquises
À Vous

**Directions:** Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller aux îles Marquises?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

**N.B.:** Only learners who did not receive corrective feedback on the preceding question saw this one.
La vallée de Haneu

Taiohae, le 19 février

Chérie,

Après 9 jours sur les mers, j’étais content de remettre les pieds sur terre. Je suis maintenant à Nuku Hiva, la plus grande et la plus peuplée des îles Marquises, où on est arrivé samedi matin. Ce jour-là, après avoir regardé les produits artisanaux aux marchés en plein air, on a vu un spectacle de danse traditionnelle marquise. Hier, on s’est arrêté pour faire un pique-nique sur le chemin de la vallée Taipivai, site archéologique rendu célèbre par l’écrivain Herman Melville. Ce matin, on a visité la vallée de Hakau, où se trouve l’une des plus hautes chutes d’eau du monde. Dans quelques heures, on va partir à Tahiti, d’où je t’écrirai encore.

Bisous,
Patrice
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Qu’est-ce que Patrice a regardé aux marchés en plein air?

- A Des fruits tropicaux.
- B Des objets d’artisanat.
- C Des T-shirts.

Keyed response: B
Vrai ou Faux?

**Directions:** Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Pendant son voyage aux îles Marquises, Patrice a lu un roman d’Herman Melville.

- Vrai
- Faux

Keyed response: Faux
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Après son séjour aux îles Marquises, où va Patrice?

Keyed response: (à) Tahiti
À Vous

*Directions:* Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller à Tahiti?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

*N.B.*: Only learners who did not receive corrective feedback on the preceding question saw this one.
Mânui a Paraīta

Papeete, le 23 février

Chérie,

Après avoir passé ces deux derniers jours ici, je dois constater que la verdure et les belles plages de Tahiti ressemblent pas mal à celles de Nuku Hiva. À notre arrivée mercredi, on a fait une randonnée dans la vallée de la Fakahoa, où se trouvent les ruines du fort éponyme, l’un des lieux clés de la guerre franco-tahitienne. Hier, on a fait le tour de la ville, surtout Mapuru a Paraīta (le Marché Municipal), l’un des meilleurs endroits pour se rendre compte de l’authenticité polynésienne. On a aussi visité l’Hôtel de Ville, une splendide réplique de l’ancien palais royal de la reine Pomaré. Aujourd’hui, on est allé au Parc Bougainville. On part bientôt aux Philippines, d’où je te réécrirai plus tard.

Bisous,
Patrice
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Qu’est-ce que Mapuru a Paraïta?

- A Un centre commercial.
- B Un musée historique.
- C Un sport traditionnel tahitien.

Keyed response: A
Vrai ou Faux?

Directions: Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

L’Hôtel de Ville de Papeete était autrefois un palais.

○ Vrai
○ Faux

Keyed response: Faux
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

À à É è ë Ï î ï Ó ô

Après son séjour à Tahiti, où va Patrice?

[Map of the world with Tahiti and the Philippines highlighted]

Keyed response: (aux) Philippines
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller aux Philippines?

☐ Absolument!

☐ Pourquoi pas?

☐ Pas vraiment.

☐ Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Le trajet entre Tahiti et les Philippines était encore plus long que celui entre la Guadeloupe et les îles Marquises: 10 jours! Cela fait 4 jours que je suis à Manille, où je m’amuse beaucoup. Mardi, on a visité Intramuros, la plus vieille partie de la ville, construite et fortifiée par les Espagnols au 16ème siècle. Mercredi, je suis allé au Luneta Park, l'un des plus grands parcs en Asie du Sud-Est. Plus tard, j'ai pris un calesa (une voiture à cheval) à Chinatown où j'ai goûté un dim sum absolument délicieux. Hier j'ai fait une promenade le long du boulevard Roxas, d'où j'ai regardé un merveilleux coucher de soleil sur la baie de Manille. Ce soir, on part au Viêt-Nam, d'où je t'écrirai ensuite.

Bisous,
Patrice
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Qu’est-ce qu’un calesa?

☐ A Un aliment.

☐ B Une danse.

☐ C Un véhicule.

Check Answer

Keyed response: C
Vrai ou Faux?

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Pendant son séjour aux Philippines, Patrice a mangé dans un restaurant chinois.

- [ ] Vrai
- [ ] Faux

Keyed response: Vrai
Après son séjour aux Philippines, où va Patrice ?

Keyed response: (au) Viêt-Nam
À Vous

*Directions:* Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller au Viêt-Nam?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Je viens de passer quelques jours intenses ici au Viêt-Nam. Dimanche, on a visité le Musée de la Guerre, une exposition poignante du point de vue vietnamien sur le conflit qui a déchiré le pays. Lundi, on a pris le car pour aller aux tunnels de Cu Chi, un réseau souterrain à 200km de long utilisé par les Viet cong pendant la guerre. Mardi, on a visité le Palais de la Réunification, ancienne résidence du président de la République Vietnamienne du Sud et lieu du dernier assaut du nord. Quelque chose de plus joyeux, hier, on est allé au marché Ben Thanh où on vend partout des produits artisanaux: laques, céramiques, marbres sculptés, vêtements en soie. Dans quelques heures, on part en Indonésie, d’où je t’écrirai de nouveau.

Tendrement,
Patrice
Vrai ou Faux?

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Les tunnels de Cu Chi sont en dessous de Hô-Chi-Minh-Ville.

- [ ] Vrai
- [ ] Faux

**Keyed response:** Faux
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Comment peut-on caractériser le séjour de Patrice au Viêt-Nam?

- Amusant.
- Ennuyeux.
- Sombre.

Keyed response: C
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

À à É è Ê ê Ë ë Ï ï Ì ì Ô ô

Après son séjour au Viêt-Nam, où va Patrice?

Keyed response: (en) Indonésie
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller en Indonésie?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Heureusement que les Indonésiens sont généralement sympathiques et chaleureux, car Jakarta est la ville la plus noire que j’aie jamais visitée. Samedi, après une longue attente pour monter au sommet du Monument National, on n’a quasiment rien vu à cause de la fumée des voitures. Dimanche, on a visité le vieux quartier de Batavia, mais il n’y reste que très peu du patrimoine des hollandais malgré trois siècles et demi de leur présence. Hier, on est allé voir le musée d’architecture "Mini Indonésie" qui expose les maisons traditionnelles et l’artisanat des 27 provinces du pays dans un grand parc au sud de la ville. Ce musée était intéressant, mais j’attends avec impatience mon départ en Inde. Je t’écrirai plus tard.

Bisous,
Patrice
Vrai ou Faux?

**Directions:** Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Patrice trouve Jakarta moins beau que Hô-Chi-Minh-Ville.

- [ ] Vrai
- [ ] Faux

**Keyed response:** Vrai
Multiple Choice

*Directions:* Click on the letter of the answer choice that best fits the question.

Pourquoi est-ce que Patrice n’a pratiquement rien vu depuis le sommet du Monument National?

- [ ] A C’est trop petit.
- [ ] B Il faisait nuit.
- [ ] C Il y avait trop de pollution.

Keyed response: C
Short Answer

**Directions:** Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

**Click on the buttons below to type the accented letters.**

À à É é È è Í í Ì ì Ò ò Ù ù Ó ó Ò ò Ù ù

Après son séjour en Indonésie, où va Patrice?

---

**Keyed response:** (en) Inde
À Vous

**Directions:** Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller en Inde?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Je viens de passer les quatre derniers jours à Pondichéry, une simple ville de taille moyenne sans trop de monuments à voir, mais intéressante en ce qu'elle est composée de deux parties. D'abord on a la ville blanche, habité à l'époque par les colonisateurs français. Les maisons y sont jolies, entourées par des murets dissimulant de jolis jardins. Les rues y sont larges, propres, arborées, on se croirait vraiment quelque part en France. La seconde partie est la ville noire, habitée par la population indienne. C'est plus étroit, les maisons sont beaucoup moins jolies, il n'y a ni arbre ni jardin. Mais c'est beaucoup plus animé aussi. Ce soir on part au Pakistan, d'où je t'écrirai de nouveau.

Bisous,
Patrice
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters:

A à É è ë ï î Ï ò

Où en Inde est-ce que Patrice vient de passer les quatre derniers jours?

Keyed response: (à) Pondichéry
À Vous

**Directions:** Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller à Pondichéry?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

**N.B.:** Only learners who did not receive corrective feedback on the preceding question saw this one.
Vrai ou Faux?

Directions: Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Pondichéry est une ancienne colonie française.

- Vrai
- Faux

Keyed response: Vrai
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

After his stay in Pondichéry, where is Patrice going?

Keyed response: (au) Pakistan
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller au Pakistan?

○ Absolument!
○ Pourquoi pas?
○ Pas vraiment.
○ Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,


Bisous,
Patrice
Vrai ou Faux?

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Karachi est une ville très animée.

○ Vrai
○ Faux

[Check Answer]

**Keyed response:** Vrai
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Comment Patrice a-t-il fait son dernier tour de la ville?

☐ a. En bus.

☐ b. En pousse-pousse.

☐ c. En taxi.

Keyed response: B
Short Answer

**Directions:** Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Click on the buttons below to type the accented letters.

À â É è Ê ê Î î Ï ï ô

Après son séjour au Pakistan, où va Patrice?

Keyed response: (à) Madagascar
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller à Madagascar?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

Les paysages de Madagascar sont les plus étonnants que j’aie jamais vus. Ils ressemblent vraiment à ceux qu’on imagine sur une autre planète. Lundi, on est allé au Parc National Tsingy de Bemaraha, site du patrimoine mondial de l’UNESCO et grande réserve naturelle où se trouvent de véritables cathédrales de calcaire, formées par un dépôt de fossiles et de coquillages morts sous la mer il y a 200 millions d’années, et par la suite façonnées par l’eau des pluies, il y a 5 millions d’années. Hier, on a fait un pique-nique au bord de la mer auprès du Cirque Rouge, amphithéâtre naturel spectaculaire aux couleurs rose et bordeaux. Demain matin, on part en Afrique de Sud, d’où je t’écrirai encore.

Bisous,

Patrick
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Qu’est-ce que Tsingy de Bemaraha?

- A. Un personnage célèbre.
- B. Une réserve nationale.
- C. Une vieille cathédrale.

Keyed response: B
Vrai ou Faux?

*Directions*: Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Patrice a vu un spectacle au Cirque Rouge.

- Vrai
- Faux

**Keyed response**: Faux
Après son séjour à Madagascar, où va Patrice?

Keyed response: (en) Afrique du Sud
À Vous

Directions: Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller en Afrique du Sud?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chère,

L'Afrique du Sud a tout. Dimanche on a jouit d'une vue spectaculaire du Cap depuis le téléphérique qui monte au sommet de la Table Mountain. Lundi, on est descendu au Cap de Bonne Espérance, la pointe australe de la péninsule du Cap où se joignent les océans Atlantique et Indien. Mardi, on est allé à Boulders Beach, où on a vu une colonie de pingouins, et hier, on a fait la route des vins de Stellenbosch. Je m'amuse tellement bien que je n'ai pas envie de partir, mais on s'en va cet après-midi au Cameroun. Je t'écrirai de là-bas.

Bisous,
Patrice
Vrai ou Faux?

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

Patrice a fait une randonnée au sommet de la Table Mountain.

- Vrai
- Faux

Keyed response: Faux
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Qu’est-ce que Stellenbosch?

☐ A Une plage.

☐ B Un vignoble.

☐ C Un zoo.

Keyed response: B
Après son séjour en Afrique du Sud, où va Patrice?

Keyed response: (au) Cameroun
À Vous

*Directions:* Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller au Cameroun?

☐ Absolument!

☐ Pourquoi pas?

☐ Pas vraiment.

☐ Pas du tout.

N.B.: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chérie,

C'est maintenant ma quatrième journée au Cameroun. Le port d'entrée de Douala n'est pas remarquable, mais j'ai bien aimé le voyage en car à Buea, le village le plus élevé sur les pistes du volcan actif du Mont Cameroun. D'abord il reste toujours des coulés de lave de l'éruption de 1999, parfois 7 à 10 mètres de hauteur, qui se dressent dans la jungle verdoyante. Ensuite, il y a un marché au bord de la route, où on peut acheter quoi que ce soit et à des prix défiant toute concurrence. Enfin, la ville elle-même témoigne toujours par son architecture charmante, son passé colonial allemand. Je prépare maintenant mon départ au Maroc, le dernier arrêt de cette odyssée. Je t'écrirai encore de là-bas.

Bisous,
Patrice
Vrai ou Faux?

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is *True* or *False.*

Patrice préfère le village de Buea au port de Douala.

○ Vrai
○ Faux

[Check Answer]

*Keyed response:* Vrai
Multiple Choice

Directions: Click on the letter of the answer choice that best fits the question.

Le Cameroun est une ancienne colonie de quel pays?

☐ A  L'Allemagne.

☐ B  L'Espagne.

☐ C  L'Italie.

Keyed response: A
Short Answer

Directions: Type a short, natural response in the box beneath the question below. Do NOT write a complete sentence. Use standard capitalization (i.e., both upper and lowercase letters) and spelling (including accents).

Après son séjour au Cameroun, où va Patrice?

Keyed response: (au) Maroc
À Vous

**Directions:** Click on your personal response to the question before continuing.

Est-ce que vous aimeriez un jour aller au Maroc?

- Absolument!
- Pourquoi pas?
- Pas vraiment.
- Pas du tout.

**N.B.**: Only learners who did not receive corrective feedback on the preceding question saw this one.
Chère,

C'est le dernier jour de ce qui a vraiment été un voyage fantastique. Casablanca, comme Pondichéry, est divisé en deux parties. La vieille ville de la Médina se caractérise par des ruelles sombres et étroites où on découvre des mosquées et des marchés. Par contre, dans la ville nouvelle, les bâtiments officiels se mêlent aux commerces le long de grandes avenues. Hier, on a visité la Mosquée Hassan II, la deuxième plus grande mosquée du monde après la Mecque, qui se trouvent en fait sur une étendue de terre asséchée au bord de la ville. Au moment où tu recevras cette carte, je serai sans doute déjà rentré en France, mais je te l'envoie pour compléter cette aventure pendant laquelle je te gardais toujours à l'esprit. Je t'embrasse très fort.

Patrice
Multiple Choice

**Directions:** Click on the letter of the answer choice that best fits the question.

Patrice compare Casablanca à quelle autre ville qu’il a visitée ?

- [ ] Jakarta.
- [ ] Karachi.
- [ ] Pondichéry.

Keyed response: C
**Vrai ou Faux?**

*Directions:* Read the following statement. Then click on the appropriate button to indicate whether the statement is True or False.

La Mosquée Hassan II se trouve dans la vieille partie de la ville.

- [ ] Vrai
- [ ] Faux

**Keyed response:** Faux
Après son séjour au Maroc, où va Patrice?

Keyed response: (en) France
À Vous

**Directions:** Click on your personal response to the question before continuing.

Est-ce que vous êtes déjà allé(e) en France?

- Non, mais j’aimerais un jour y aller.
- Non, et je n’ai pas vraiment envie d’y aller.
- Oui, et je m’y suis bien amusé(e).
- Oui, mais ce n’était rien de spécial.

**N.B.:** Only learners who did not receive corrective feedback on the preceding question saw this one.
Congratulations!

You have now finished the study portion of today's lesson. We now ask that you take a few moments to complete a short survey to give us some feedback on your experience. Please click 'Start Survey' to begin.
Chérie,

Ça ne fait même pas deux heures depuis qu’on s’est parlé au téléphone mais tu me manques déjà. Je souhaiterais que tu puisses m’accompagner dans ce voyage et vivre cette aventure avec moi. Je promets de t’envoyer des nouvelles de chaque pays. J’arrive aux États-Unis mercredi. Je t’écrirai encore vers la fin de mon séjour là-bas. Maintenant je dois finir de faire mes bagages donc je te laisse pour l’instant. Je t’embrasse très fort.

Patrice
New York, le 3 février 2006

Chérie,

Le voyage transatlantique était plutôt agréable. On est arrivé à New York mercredi matin. Ce jour-là, après un tour du siège de l'ONU, on a visité les quartiers de Manhattan, y compris Chinatown, Little Italy et Spanish Harlem. Les États-Unis, c'est un pays vraiment diversifié! Jeudi on a visité Ellis Island et la statue de la Liberté. Hier on a visité le musée Guggenheim dans la journée et on est allé voir «Le Roi Lion» à Broadway le soir. Notre séjour ici se termine, on part bientôt aux Petites Antilles, d'où je t'écrirai de nouveau.

Tendrement,
Patrice
Chère,

Alors qu'il ne faisait que 1° à New York, ici il fait presque 25°. C'est formidable! La Guadeloupe est vraiment belle. Mardi on a visité le Fort Louis Delgrès et la Cathédrale de Basse-Terre, deux beaux exemples du patrimoine architectural français datant respectivement du 17ème et 18ème siècle. Hier, on a fait une randonnée au sommet du volcan de la Soufrière, le point culminant de l'archipel. Ce matin, on a visité le Jardin Botanique de Deshaies, un espace merveilleux appartenant autrefois au comédien Coluche. Ce soir on commence un long voyage aux îles Marquises en passant par le canal de Panama. Je t'écrirai davantage de là-bas.

Tendrement,
Patrice
Chérie,

Après 9 jours sur les mers, j’étais content de remettre les pieds sur terre. Je suis maintenant à Nuku Hiva, la plus grande et la plus peuplée des îles Marquises, où on est arrivé samedi matin. Ce jour-là, après avoir regardé les produits artisanaux aux marchés en plein air, on a vu un spectacle de danse traditionnelle marquise. Hier, on s’est arrêté pour faire un pique-nique sur le chemin de la vallée Taipivai, site archéologique rendu célèbre par l’écrivain Herman Melville. Ce matin, on a visité la vallée de Hanei, où se trouve l’une des plus hautes chutes d’eau du monde. Dans quelques heures, on va partir à Tahiti, d’où je t’écrirai encore.

Bisous,
Patrice
Chérie,

Après avoir passé ces deux derniers jours ici, je dois constater que la verdure et les belles plages de Tahiti ressemblent pas mal à celles de Nuku Hiva. À notre arrivée mercredi, on a fait une randonnée dans la vallée de la Fachoda, où se trouvent les ruines du fort éponyme, l'un des lieux clés de la guerre franco-tahitienne. Hier, on a fait le tour de la ville, surtout Mapuru a Paraîta (le Marché Municipal), l'un des meilleurs endroits pour se rendre compte de l'authenticité polynésienne. On a aussi visité l'Hôtel de Ville, une splendide réplique de l'ancien palais royal de la reine Pomaré. Aujourd'hui, on est allé au Parc Bougainville. On part bientôt aux Philippines, d'où je te réécrirai plus tard.

Bisous,
Patrice
Chérie,

Le trajet entre Tahiti et les Philippines était encore plus long que celui entre la Guadeloupe et les îles Marquises: 10 jours! Cela fait 4 jours que je suis à Manille, où je m’amuse beaucoup. Mardi, on a visité Intramuros, la plus vieille partie de la ville, construite et fortifiée par les Espagnols au 16e siècle. Mercredi, je suis allé au Luneta Park, l’un des plus grands parcs en Asie du Sud-Est. Plus tard, j’ai pris un calèche (une voiture à cheval) à Chinatown où j’ai goûté un dim sum absolument délicieux. Hier j’ai fait une promenade le long du boulevard Roxas, d’où j’ai regardé un merveilleux coucher de soleil sur la baie de Manille. Ce soir, on part au Viêt-Nam, d’où je t’écrirai ensuite.

Bisous,
Patrice
Chérie,

Je viens de passer quelques jours intenses ici au Viêt-Nam. Dimanche, on a visité le Musée de la Guerre, une exposition poignante du point de vue vietnamien sur le conflit qui a déchiré le pays. Lundi, on a pris le car pour aller aux tunnels de Cu Chi, un réseau souterrain à 200km de long utilisé par les Viet cong pendant la guerre. Mardi, on a visité le Palais de la Réunification, ancienne résidence du président de la République Vietnamienne du Sud et lieu du dernier assaut du nord. Quelque chose de plus joyeux, hier, on est allé au marché Ben Thanh où on vend partout des produits artisanaux: laques, céramiques, marbres sculptés, vêtements en soie. Dans quelques heures, on part en Indonésie, d’où je t’écrirai de nouveau.

Tendrement,

Patrice
Chérie,

Heureusement que les Indonésiens sont généralement sympathiques et chaleureux, car Jakarta est la ville la plus noire que j'aie jamais visitée. Samedi, après une longue attente pour monter au sommet du Monument National, on n’a quasiment rien vu à cause de la fumée des voitures. Dimanche, on a visité le vieux quartier de Batavia, mais il n’y reste que très peu du patrimoine des hollandais malgré trois siècles et demi de leur présence. Hier, on est allé voir le musée d'architecture "Mini Indonésie" qui expose les maisons traditionnelles et l'artisanat des 27 provinces du pays dans un grand parc au sud de la ville. Ce musée était intéressant, mais j’attends avec impatience mon départ en Inde. Je t’écrirai plus tard.

Bisous,
Patrice
La ville blanche de Pondichéry

Pondichéry, le 28 mars

Chère,

Je viens de passer les quatre derniers jours à Pondichéry, une simple ville de taille moyenne sans trop de monuments à voir, mais intéressante en ce qu'elle est composée de deux parties. D'abord on a la ville blanche, habitée à l'époque par les colonisateurs français. Les maisons y sont jolies, entourées par des murets dissimulant de jolis jardins. Les rues y sont larges, propres, arborées, on se croirait vraiment quelque part en France. La seconde partie est la ville noire, habitée par la population indienne. C'est plus étroit, les maisons sont beaucoup moins jolies, il n'y a ni arbre ni jardin. Mais c'est beaucoup plus animé aussi. Ce soir on part au Pakistan, d'où je t'écrirai de nouveau.

Bisous,
Patrice
Ché ric,

Dans la plus grande ville du Pakistan les vieux immeubles côtoient les bâtiments flamants neufs. Samedi, on a visité Wazir Mansion, lieu de naissance de Muhammad Ali Jinnah, fondateur du pays, ainsi que le Mausolée Quaid-i-Azam, lieu de son enterrement. Dimanche, on est allé au Musée Mohatta, ancienne résidence palatiale de sa sœur, Fatimah Jinnah. Lundi, j'ai visité Frere Hall, témoin du Raj britannique et l'une des plus belles structures de la ville. Autrefois l'Hôtel de Ville, il renferme actuellement une grande bibliothèque et une galerie d'art. Hier, j'ai fait un dernier tour de la ville en rickshaw car on part aujourd'hui à Madagascar. Je t'écrirai de nouveau dans quelques jours.

Disous,
Patrice
Chérie,

Les paysages de Madagascar sont les plus étonnants que j’aie jamais vus. Ils ressemblent vraiment à ceux qu’on imagine sur une autre planète. Lundi, on est allé au Parc National Tsingy de Bemaraha, site du patrimoine mondial de l’UNESCO et grande réserve naturelle où se trouvent de véritables cathédrales de calcaire, formées par un dépôt de fossiles et de coquillages morts sous la mer il y a 200 millions d’années, et par la suite façonnées par l’eau des pluies, il y a 5 millions d’années. Hier, on a fait un pique-nique au bord de la mer auprès du Cirque Rouge, amphithéâtre naturel spectaculaire aux couleurs rose et bordeaux. Demain matin, on part en Afrique de Sud, d’où je t’écrirai encore.

Bisous,
Patrice
Chérie,

L'Afrique du Sud a tout. Dimanche on a jouit d'une vue spectaculaire du Cap depuis le téléphérique qui monte au sommité de la Table Mountain. Lundi, on est descendu au Cap de Bonne Espérance, la pointe australe de la péninsule du Cap où se joignent les océans Atlantique et Indien. Mardi, on est allé à Boulders Beach, où on a vu une colonie de pingouins, et hier, on a fait la route des vins de Stellenbosch. Je m'amuse tellement bien que je n'ai pas envie de partir, mais on s'en va cet après-midi au Cameroun. Je t'écrirai de là-bas.

Bisous,
Patrice
Chérie,

C'est maintenant ma quatrième journée au Cameroun. Le port d'entrée de Douala n'est pas remarquable, mais j'ai bien aimé le voyage en car à Buea, le village le plus élevé sur les pistes du volcan actif du Mont Cameroun. D'abord il reste toujours des coulées de lave de l'éruption de 1999, parfois 7 à 10 mètres de hauteur, qui se dressent dans la jungle verdoyante. Ensuite, il y a un marché au bord de la route, où on peut acheter quoi que ce soit et à des prix défiant toute concurrence. Enfin, la ville elle-même témoigne toujours par son architecture charmante, son passé colonial allemand. Je prépare maintenant mon départ au Maroc, le dernier arrêt de cette odyssée. Je t'écrirai encore de là-bas.

Bisous,
Patrice
Chérie,

C’est le dernier jour de ce qui a vraiment été un voyage fantastique. Casablanca, comme Pondichéry, est divisé en deux parties. La vieille ville de la Médina se caractérise par des ruelles sombres et étroites où on découvre des mosquées et des marchés. Par contre, dans la ville nouvelle, les bâtiments officiels se mêlent aux commerces le long de grandes avenues. Hier, on a visité la Mosquée Hassan II, la deuxième plus grande mosquée du monde après la Mecque, qui se trouvent en fait sur une étendue de terre asséchée au bord de la ville. Au moment où tu recevras cette carte, je serai sans doute déjà rentré en France, mais je te l’envoie pour compléter cette aventure pendant laquelle je te gardais toujours à l’esprit. Je t’embrasse très fort.

Patrice
APPENDIX I: Post-treatment Survey

Survey Question #1 (of 8)

Think about similar activities you have done that relied primarily on textbooks and paper-and-pencil exercises. Compared to those activities:

Because of the way this activity presented the material in an interactive computer-based format, how enjoyable was the learning process?

☐ Much More Enjoyable
☐ Somewhat More Enjoyable
☐ About as Enjoyable
☐ Somewhat Less Enjoyable
☐ Much Less Enjoyable

Continue
Survey Question #2 (of 8)

Indicate how strongly you agree or disagree with the following statement:

Because of the way this activity presented the material in an interactive computer-based format, I was better able to understand the ideas and concepts taught in it.

- [ ] Strongly Agree
- [ ] Agree
- [ ] Disagree
- [ ] Strongly Disagree

Continue
Survey Question #3 (of 8)

Indicate how strongly you agree or disagree with the following statement:

Because of the way this activity presented the material in an interactive computer-based format, I was at a disadvantage because I do not possess adequate computer skills.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree
Survey Question #4 (of 8)

Indicate how strongly you agree or disagree with the following statement:

Because of the way this activity presented the material in an interactive computer-based format, I was at a disadvantage because I do not possess adequate typing skills.

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

Continue
Survey Question #5 (of 8)

Indicate how strongly you agree or disagree with the following statement:

I’d be interested in trying out similar experimental interactive computer-based French materials outside of class.

- ☐ Strongly Agree
- ☐ Agree
- ☐ Disagree
- ☐ Strongly Disagree

[Continue]
Survey Question #6 (of 8)

Indicate how strongly you agree or disagree with the following statement:

I’d be interested in participating in UA French Department research studies outside of class even if the proposed activities did not involve interactive computer software.

- [ ] Strongly Agree
- [ ] Agree
- [ ] Disagree
- [ ] Strongly Disagree

[Continue]
Survey Question #7 (of 8)

How would you characterize your French 102 course this semester?

☐ part of my major
☐ part of my minor
☐ a Gen Ed requirement
☐ an elective
☐ other

[Continue]
Survey Question #8 (of 8)

How old will you be at the end of this course (i.e., on Dec. 11, 2006)?

☐ under 18
☐ 18 - 20
☐ 21 - 23
☐ over 23
Merci Beaucoup!

This is the end of the survey. Please click 'QUIT' to exit the program.
APPENDIX J: Human Subjects Recruitment Script

“As students in FREN 102 this semester, you were given the opportunity to try out some specially developed computer-based reading materials. You also took a number of cloze tests to help your instructor identify your linguistic strengths, address your weaknesses, and monitor your progress. Now, I would like to ask permission to use your coursework to help shape not only the UA French program, but other language programs, both at UA and other institutions. Specifically, we want to examine your answers to the two ungraded practice cloze tests, the extra credit diagnostic test, and the two graded reading assessments to uncover changes in patterns of difficulty over time. We would also look at the log of your activities during the computerized reading lesson to determine their effects on your subsequent reading performance. Finally, we would use your responses to the post-lesson survey to help us decide whether and how to develop and implement this type of technology in the future.

“Your identity will be kept completely confidential, and whether or not you decide to permit the use of your data for this project will not affect your grade. In fact, your instructor will not even know of your decision. However, federal regulations require those of us in charge of the project to obtain consent to use your data for any non-departmental purposes, such as publication or formal presentation. We hope that all of you will agree to participate so that our final results will be as representative as possible. This is a chance for you to have an impact on how French and other languages are taught at the U of A and possibly elsewhere. We now ask that you kindly grant us permission to use your data by signing the consent form we have prepared for you.”

Note to Recruiter: Next distribute the consent forms and read them to the students point by point. After you ask them to sign, tell them that if they would like a personal copy of the form they should circle the word ‘copy’ on the first page, in the introduction. They can then pick up their copy from me in ML 301 either just before or during the exam period (i.e., 8-10PM, Mon., 12/11). Otherwise, if they would prefer, they can add their email address below their signature, and they will receive a pdf copy as an attachment.
REFERENCES


