Feedback in Intelligent Computer-Assisted Language Learning and Second Language Acquisition: A study of its effect on the acquisition of French past tense aspect using an Intelligent Language Tutoring System

Ruth Mary Hanson, B.A.(Hons), M.A.

A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy

to
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

January 2007

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Ruth Mary Hanson

Abstract

Questions surrounding the impact of feedback in response to learner error are of interest in the fields of both Second Language Acquisition (SLA) and Intelligent Computer-Assisted Language Learning (ICALL). Current empirical SLA research seeks to ascertain what feedback types have a statistically significant positive impact on the process of acquiring a second language. Similarly, research in ICALL focuses on testing Intelligent Language Tutoring Systems (ILTSs) generally as well as the effectiveness of the feedback that they deliver. Despite this common interest in feedback, to date there has been no significant interdisciplinary research involving the two fields.

The experiment reported here seeks to bridge this gap. Using a purpose-built ILTS, we tested the effect of two types of feedback on the acquisition of French past tense aspect among anglophone learners. Inspired by previous work in SLA, Explicit Inductive (EI) and Input Processing (IP) feedback were tested against a control group using a pre test/post test design. The learners completed a transformation and a grammaticality judgment task. For the transformation, they were presented with texts in the present tense and asked to re-write them in the past tense. For the grammaticality judgment, they had to rank the grammaticality of each sentence in a set of texts. In response to errors, EI feedback interpreted the aspectual meaning of the learners' answer and explicitly told them that it was not the most natural according to the context. In order to encourage form-meaning mapping, IP feedback asked the learners to match their erroneous answer to its interpretation. Two interpretations were presented: one was the target interpretation and the other matched the learner's answer. Having made their choice, they were then told whether it was correct as well as which interpretation was in fact target-like.

The quantitative evaluation of the effectiveness of the EI and IP feedback was not statistically significant. We argued that this was due to a combined effect of learner level, target structure and feedback.

Acknowledgements

To my supervisors. Thank you to Professor Antonella Sorace for encouraging me to write less tersely and to Dr. Helen Pain for a lot of support and encouragement.

To various sources of financial support. The opportunity for me to do my PhD at the University of Edinburgh was made possible by generous funding from the College of Humanities and Social Sciences and the Overseas Research Students Award Scheme. Thank you to Colin Coates, who was the Director of Canadian Studies when I arrived, for providing me with two bursaries courtesy of the Canadian High Commission.

To a couple of people who gave me indispensable help. Thank you to Peter Dayan for administrative help in procuring French first and second language speakers. I am truly indebted to Ruli Manurung for solving the coding problems in the tutoring system that were beyond me and for making me feel a lot less incompetent as a programmer.

To my subjects. There would not have been an experiment to conduct, if it hadn't been for them. To the (poor, unsuspecting) anglophones, thank you for enduring what must certainly have been a laborious, and possibly traumatic, experience for a mere tenner. Aux francophones, et à Mathieu Joron en particulier, merci de votre aide et de votre gentillesse. J'aimerais aussi remercier Alice Foucart d'avoir toujours été prête à m'aider à comprendre les subtilités de sa langue maternelle.

To my friends. A great big bear hug of a thank you to Cassie and Rob for all manner of kindnesses: giving me a job as a data collector, printing and submitting my

thesis and advising me about statistics. To Marisa for painstakingly proofreading parts of the thesis and for much hilarity, particularly when we shared an office. To Ivan for his warmth, generosity of spirit and great cooking. To Susana for being a terrific friend and always laughing whole-heartedly at my jokes. To Christine for great chats and commiseration.

To my extended and immediate family. Having relatives only a one-hour train ride away was always a great source of happiness and reassurance to me. Thank you to AA and UD for regular phone calls and cosy visits and to Celia and Tom for their constant support and generosity to an impecunious student. To Jen for being the world's best correspondent and to Fiona for listening. To my parents for giving me strength and to my dad for providing technical support and actually reading the thesis. The four of you have been immortalized in Example 3.4!

To Bill for being an endless source of understanding, encouragement and joy.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Input, Positive Evidence and Negative Evidence in Second Language Acquisition

1.1.1 Input

Input is a term in the second language acquisition (SLA) literature which is generally understood to refer to written or spoken manifestations of the second language which are available to the learner. It has been distinguished from intake, as far back as Corder (1967). While he defines input as "what is *available* for going in" to the mind of the learner, intake is "what goes in" (Corder 1967, p. 165, emphasis in the original). In other words, what the learner is exposed to in the external second language environment (i.e. input) is distinct from what he or she extracts from it (i.e. intake).

More recent research, which has addressed the issue of input, exhibits divergent views on what input is and how it might fit into a model of second language acquisition. For instance, Krashen (1982) talks about input as being comprehensible to, or comprehended by, the learner. According to his *Comprehensible Input Hypothesis*, one of the conditions for acquisition to proceed is that input which learners are able to comprehend contain "structures a bit beyond the acquirer's current level" (Krashen 1982, p. 33). Carroll (2001), on the other hand, discusses input within a framework of speech processing. She makes a distinction between input to speech parsers and input to learning mechanisms. The

former are structured representations of transduced linguistic stimuli which are processed by speech parsers. The latter is information derivable from analyses or inferences from the parsing of speech stimuli (positive evidence) or information derivable from strings which are not part of the second language grammar (negative evidence). It seems that Krashen's view of input does not distinguish between what is internal and external to the learner. If input is comprehensible or comprehended, then it must entail some kind of representation in the mind of the learner. However, the idea that input might contain structures still to be acquired seems to indicate that it can also be understood to be external to the learner. In contrast, both types of input identified by Carroll clearly reside inside the mind of the learner. In fact, she uses a separate term to refer to information which enters the learner's perceptual system of the second language environment; namely, *stimuli*. For further work which also addresses these issues, see Schwartz (1986, 1987, 1993) and Long (1981).

1.1.2 Positive Evidence

Part of the explanation of how input comes to be represented in the mind of the learner involves elaboration of the role of positive and negative evidence. Pertaining to first language acquisition, Pinker (1989) defines positive evidence as evidence about what word strings are grammatical in the target language which, in turn, informs the child that his or her hypothesis about the target language is wrong. The same definition applies in second language acquisition. Thus, when learners parse strings in their second language, they formulate analyses which give them information about what is permissable in the second language grammar. For instance, from parsing of the string *On va souvent au cinéma* (We often go to the cinema), an anglophone learner can derive evidence for adverb placement in French; namely, that adverb phrases follow verb phrases.

1.1.3 Negative Evidence

Negative evidence is information about what is not part of the target-language grammar. Pinker (1989) raises a number of theoretical conditions which would have to be met in order for negative evidence to play a role in first language acquisition. We will discuss each point in turn, as it relates to second language acquisition.

The first question is whether or not negative evidence exists for children. One of Pinker's reasons against the existence of negative evidence is that parental feedback does not occur on all ungrammatical sentences for all children. This is also true in second language acquisition. Documented evidence from both classroom settings (Allwright 1975, Courchene 1980, Long 1988, Lyster & Ranta 1997, Swain & Carroll 1987) and naturalistic interactions between native speakers and second language learners (Chun et al. 1982, Crookes & Rulon 1988, Day et al. 1984, Gaskill 1980) suggests that a large proportion of errors go uncorrected. For instance, in their database of 18.3 hours of recorded communicative interaction in a French immersion classroom setting, Lyster & Ranta (1997) report that 62% of student turns containing at least one error or use of the L1 were addressed with feedback of various sorts from the teacher. This left a sizable 38% of errors in student turns which were not corrected. We believe that Intellient Computer-Assisted Language Learning (ICALL) offers a way of making negative evidence available to learners through the use of Intelligent Language Tutoring Systems (ILTSs). Exploiting artificial intelligence techniques (for more details see Section 2.2.1), ILTSs have the advantage of providing informative, individualized, consistent feedback on all ungrammatical sentences in a given domain.

Assuming that negative evidence existed for children, Pinker's next question is whether it would be useful. He argues that because there are many rules and principles required to generate a sentence, it may not be possible for children to identify the source of an error from negative evidence. He calls this the "blame-assignment" problem. We argue, in Chapter 5, that one of the benefits of ICALL for SLA is that ILTSs can constrain the language learning task and provide informative feedback such that learners should be able to identify the source of their errors within a given area of grammar (for further discussion of this point, see Section 5.3).

The type of feedback provided to the learner is also a factor in the usefulness of negative evidence. A number of studies, which will be reviewed in detail in Section 2.3, have tested a wide range of feedback, including for instance recasts, clarification requests, models and metalinguistic information. Overall, it is unclear what types of feedback might be useful to second language learners from the mixed results of studies that have been conducted to date. Further research on the question is clearly needed.

Pinker's next question is whether or not children would actually use negative evidence, assuming it were available and useful. It is not guaranteed that second language learners will use feedback either. If they do not, we must explain why this is so.

Finally, in order to claim that negative evidence has a role in explaining first language acquisition, according to Pinker, research must demonstrate that learning is due only to negative evidence, and not some other learning mechanism (e.g. positive evidence). For successful learning to occur, the learner must be able to come up with a correct hypothesis and be able to falsify any incorrect ones. Falsification of incorrect hypotheses will occur on the basis of positive evidence, except if the learner's interlanguage grammar generates a superset of the target language. Pinker raises double object verbs in English as an example of an area of grammar for which the learner will logically generate an overgeneralized hypothesis in the absence of negative evidence[1]. Consider the following sentences and their argument structure:

(1.1)   *The woman gave a present to the boy* (V+NP+PP)
           The woman gave the boy a present (V+NP+NP)

On the basis of input such as this, a learner could logically hypothesize that any verb with the argument structure V+NP+PP also allows the structure V+NP+NP. However, this is not the case, as the following examples show:

(1.2)   *The witness reported the crime to the police* (V+NP+PP)
           *The witness reported the police the crime (V+NP+NP)
(1.3)   *A benefactor donated the money to the library* (V+NP+PP)
           *A benefactor donated the library the money (V+NP+NP)

Yet, since there is nothing in the input to contradict the faulty hypothesis, it would seem that negative evidence is necessary for successful learning to take place. Negative evidence is not useful, in the sense of being necessary, for learning all areas of grammar, because in many cases, positive evidence alone will suffice. However, in these cases, negative evidence may serve to expedite the

---

[1]Although double object verbs have been studied at length, the passive, the lexical causative alternation and the locative alternation are phenomena which exhibit the same learning problem. See Pinker (1984, 1989) for explanation and examples.

acquisition process in areas of grammar which might otherwise take a very long time to acquire.

1.2 Feedback in ICALL and SLA

Despite being a topic of interest in the fields of both Intelligent Computer-Assisted Language Learning (ICALL) and Second Language Acquisition (SLA), previous research on feedback[2] has been separate. This could be because the interests of researchers in both ICALL and SLA have not converged on the issue of feedback. Most researchers in ICALL have been more interested in technical innovations or in devising language learning tasks, pedagogically motivated or otherwise. There has been a small amount of work which has acknowledged a role for feedback. However, it has not appealed to previous research in SLA for motivation, incorporating instead the intuitions of the system designers. As for the SLA literature, the relevance of computerized feedback delivery is perhaps not apparent, since valid experiments can be conducted without it. Moreover, feedback consistent with the Interaction Hypothesis (Long 1981, 1983, 1996), a framework which has dominated current experimental research on feedback in the SLA literature, does not readily lend itself to delivery via a computational medium.

Another reason for the absence of interdisciplinary research between ICALL and SLA may lie in the perception that neither field has much to gain from the knowledge amassed to date by the other. ICALL research has been dominated by the development of prototypes, few of which have attained a sufficient degree of functionality to undergo any empirical testing. Therefore, the effectiveness of Intelligent Language Tutoring Systems (ILTSs) themselves as well as their feedback has yet to be shown. Feedback has received a fair amount of attention in the SLA literature. However, the results have been so mixed that no clear conclusions have been reached about what types of feedback are most helpful to learners.

---

[2]The term feedback is defined as a response to an utterance that deviates in some way from the norms of a given target language. Depending on the field, it is also referred to as correction or negative evidence.

1.3 The Study

We believe that work on feedback has the potential to be an extremely fruitful area of interdisciplinary study that could answer questions in both ICALL and SLA research. Both fields would benefit from further empirical research on the effectiveness of feedback. For ICALL, an interdisciplinary approach would represent an opportunity to carry out empirical testing of ILTS-delivered feedback which is motivated by previous research in SLA. For SLA, it would enable further testing of feedback types that are not tied to the Interaction Hypothesis. Moreover, there would be the added methodological advantage of providing feedback from an ILTS, as it would allow for strict control over the content and amount of feedback provided to the learners (Hulstijn 1997).

With these ideas in mind, we report on a study which has integrated dimensions of feedback that are significant to both fields. The study entailed conducting an experiment whose purpose was to ascertain whether feedback would incite improvement in the acquisition of a particular area of grammar among second language learners. The feedback in question was delivered by an ILTS designed and programmed for the experiment by the author.

Two types of feedback were considered: Input Processing (IP) and Explicit Inductive (EI), both of which were inspired by previous research in the SLA literature on feedback. IP feedback follows work by Bill VanPatten and his colleagues on a type of instruction that encourages target-like mapping of meaning to form. EI feedback alerts learners to their errors and explains the meaning of a target structure on a case by case basis.

1.4 The Target Structure: Past Tense Aspect in French

Constrained by syntactic, semantic and pragmatic factors, aspect is a challenging dimension of second language acquisition, particularly for learners whose first language aspectual system is very different from that of their second language. French, the target language for the study, possesses an aspectual system which is different from other languages in that the correspondence between aspectual form and aspectual meaning is symmetrical. It is renowned for being extremely

difficult to acquire for anglophone learners, the target population of the experiment. As such, it represented a good candidate for a target structure.

1.5 Structure of the Thesis

The thesis is structured as follows:

- Chapter 2 will review the literature on feedback in both ICALL and SLA with a view to illuminating what further research is needed. This will allow us to establish the precise ways in which an interdisciplinary approach might benefit both fields as well as to motivate IP and EI feedback as worthwhile candidates for testing.
- Following a brief presentation of the theoretical terminology and underlying concepts of aspect, chapter 3 will contrast the aspectual systems of French and English for the purpose of identifying potential areas of difficulty for anglophone learners of French due to language transfer.
- Chapter 4 will review the research conducted on the acquisition of aspect by second language learners in order to motivate the choice of aspect as the target structure for the experiment. It will reveal what is already known, as well as what further work is needed, concerning how second language aspect is acquired. Furthermore, it will identify what work, if any, has explored the effect of feedback on the acquisition of aspect.
- Chapter 5 will present the methodology of the experiment. It will also contain the statistical analysis and results.
- Chapter 6 will discuss the results of the experiment, make suggestions for future research and summarize the contributions of the thesis.

CHAPTER 2

Feedback in Intelligent Computer-Assisted Language Learning and Second Language Acquisition

2.1 Introduction

The purpose of this chapter is to review what work has been carried out in ICALL and SLA relating to feedback in order to identify areas in both fields where further work is needed. It will also use these literatures to motivate and elaborate the two types of feedback we tested in our experiment. Sections 2.2 and 2.3 will review the ICALL and SLA literatures on feedback. Section 2.4 will be devoted to presenting the work that is most relevant to ours, namely that which has attempted to evaluate the effectiveness of feedback in an ILTS environment. Section 2.5 will present the work which inspired IP feedback and Section 2.6 will summarize the chapter.

2.2 Previous Work in Intelligent Computer-Assisted Language Learning (ICALL)

Intelligent Computer-Assisted Language Learning is a branch of CALL research[1] which is defined by the fact that it exploits artificial intelligence (AI) techniques

---

[1]A review of the CALL literature is beyond the scope of this thesis. We refer the interested reader to two reviews: Debski (2003) explores the directions in which CALL has moved, covering the literature from 1980 to 2000, and Liu et al. (2003) looks at the use of technology to enhance

in the creation of software for second language learning. Since the first Intelligent Language Tutoring Systems began to appear in the early 1980s, they have incorporated a range of methods for simulating intelligence, including Natural Language Processing (NLP), Natural Language Generation (NLG), User Modelling (UM) and, most recently, Automatic Speech Recognition (ASR). They have also targeted many areas of language, among them grammar (i.e. a combination of morphology and syntax, as well as semantics to a limited degree), vocabulary, orthography, stylistics, pronunciation and pragmatics. Only the ILTSs developed for tutoring grammar will be reviewed here, as they are most closely related to the programme created for this thesis. For further details of ILTSs incorporating all the areas of language and AI techniques mentioned above, the reader is referred to Bailin (1995), a bibliography of ICALL applications, and to Gamper & Knapp (2002), a review of ILTSs developed from 1993 to 2002.

2.2.1 Intelligent Language Tutoring Systems (ILTSs) for Grammar

NLP-focused ILTSs

From a pedagogical perspective, the obvious appeal of ILTSs is that they can give feedback that locates and explains learner errors. However, while the vast majority of systems have this capability, the design of the feedback is, in many cases, not an aspect of the ILTS which is given much attention. In the early work, feedback considerations tended to be overshadowed by NLP innovations, as the five examples in the following paragraphs will illustrate.

Schwind (1990) was primarily interested in designing a grammar knowledge base of German which was extensive in coverage and could be adapted to perform several functions, including sentence production, sentence analysis, error analysis and explanation and querying. While the feedback incorporated in her ILTS went beyond a simple statement explaining the error, by the author's own admission, the design of the feedback was not theoretically motivated, nor was it the aim of the research.

---

language learning for the period of 1990 to 2000. See also Jung (2005) for an extensive CALL bibliography.

Scripsi, an ILTS developed by Catt & Hirst (1990), did incorporate the phenomena of transfer and overgeneralization identified in the Second Language Acquisition (SLA) literature. However, they served to motivate an NLP technique, unique at the time, whereby parsing the learner input involved the L1 grammar as well as the L2. The feedback was a single, sometimes cryptic, statement of a violated L2 rule or an overgeneralization, which shows that the authors' concerns about the feedback did not go beyond its content. A subsequent ILTS by Wang & Garigliano (1992) used the same technique of treating transfer errors by analyzing input using a model of both the L1 and the L2 grammars with the additional innovation that they exploited a corpus of learner transfer errors made by native speakers of English learning Chinese. The authors mention that their ILTS has four types of tutorial strategy, namely coaching, correction, confirmation and revision, but do not explain how they work nor give any details on their content.

SWIM (Zock et al. 1989) was designed using explicit dialogue for the purpose of documenting, and subsequently analyzing, how learners reduce their hypothesis space in language learning, rather than in remediating particular errors. As such, the learners were required to interact with the system using four possible questions to which the system was capable of replying on the topic of personal pronouns in French. Although the authors mention that the data they collect using SWIM could be useful to psycholinguists and language educators, their main motivation seemed to be in creating a computational model of how humans learn languages for use in research on machine learning of natural languages.

The significant feature of ET (English Tutor (Fum et al. 1988)) is its capacity to generate specific hypotheses about learners' knowledge of the English verb system and to ascertain any misconceptions they may have. To achieve this, ET begins by identifying the correct answer for a given question, along with the appropriate rules of usage. It then compares this information to the learner input and begins to infer what usage rules the learner does and does not know, based on their responses to a series of questions. Once the system is able to come to a conclusion about the learner's knowledge, it provides feedback which states the rules that it thinks the learner knows as well as his or her misconceptions. As in the case of *Scripsi* (Catt & Hirst 1990) mentioned above, the fact that the

feedback is simply stated to the learner without any justification indicates that the feedback was not a priority.

Pedagogically-focused ILTSs

Subsequent work has undergone a shift towards focusing on the pedagogical considerations of designing ILTSs. This has resulted in systems whose design has been informed by a specific approach to the teaching of a given area of grammar or by particular language teaching and learning theories. As the examples of five such systems reviewed below will reveal, the pedagogical focus tends to be on ILTS tasks with little, or no, attention to feedback.

MILT (Military Language Tutor) developed by Kaplan & Holland (1995)[2] is grounded in a unitary theory of language learning and adopts a communicative approach to language teaching. As such, MILT provides a microworld in which the learners have to understand, and be understood by, an agent in order to achieve certain pre-defined goals. NLP techniques are used to interpret learner production and detect errors. In keeping with the communicative language teaching objective of fluency over accuracy, the learners are not directly informed of their errors. Instead, the feedback is implicit in that if their input is comprehensible, the agent will respond accordingly, whereas if it is not, the agent will say that he does not understand.

Collentine (1998) advocates the use of mind-centred cognitive learning theories, basing the design of his ILTS, *Subjunctive Discoveries*, on the Input Hypothesis (Krashen 1982). Drawing on the Processing Instruction framework (VanPatten 1993) and work on attention by Cowan (1995), the tasks in *Subjunctive Discoveries* are designed to provide structured input that directs the learner's attentional resources to detecting the grammatical stimulus. In practical terms, this means that learners are presented with two different situations both in video clip and text form to which they have to associate a sentence in the subjunctive in Spanish. The subjunctive elements of the sentence are made physically salient by highlighting the text in colour as well as by presenting the text aurally. To complete the task, the learners drag and drop the sentence into a space below the situation

---

[2]Holland et al. (1999) describes a more recent version of the same ILTS that uses speech recognition technology.

they think is correct. In the case of an error on the part of the learner, the feedback is not only implicit, but also non-linguistic. Rather than asking the learners if they think their choice is correct or telling them that it is not, *Subjunctive Discoveries* will simply prevent the sentence from being dropped below the incorrect situation.

Rather than being informed by a particular learning or teaching theory, Yoshii & Milne (1998) adopt their own idea of teaching article usage in English by way of diagrammatic reasoning into the design of their ILTS, DaRT (Diagrammatic Reasoning Tool). Some consideration was given to the feedback incorporated into the system in that rather than just explaining why the learner's choice of article is wrong, it provides hints about what the correct choice should be. However, this approach represents one of several features of the ILTS rather than one of the three main objectives of the research. Furthermore, it is simply described without any justification. The same criticism can be made of the *Fawlty Article Tutor* (Kurup et al. 1992). The idea of teaching the rules for article usage in English based on noun classifications constitutes the main innovation of the system and, while the feedback explains why the rule a learner selects is incorrect, this is simply described as a part of the architecture of the system and is not subject to any justification, theoretical or otherwise.

Feedback-centred ILTSs

There are a number of ILTSs whose main goal is to provide feedback that is tailored to the needs of the individual learner. Through the review of these ILTSs, presented below, a variety of approaches to the provision of feedback in ICALL will be identified.

The most frequent approach has been to use several levels of feedback that give increasingly specific information about a given error each time it is made. In *Textana* (Schulze 1999, Schulze & Hamel 2000), a prototype ILTS for learners of German, five levels of feedback are proposed. When an error is first made, *Textana* will give a warning that the sentence contains a grammatical error. If the same error is made again, *Textana* will identify the constituent where the error is located. If the error is made a third time, a general explanation of the error will be given. On the fourth attempt, *Textana* will explain the error making specific

reference to the word or words where the problem lies. Once all of these possibilities have been exhausted, *Textana* asks a leading question contrasting the incorrect answer given by the learner with the correct one.

L'haire (2004) describes a similar potential feedback scheme for *FreeText*[3] but proposes three levels of specificity. Rather than beginning with a warning about there being an error in the sentence, the first response to an error is to indicate its location (i.e. within a particular NP or VP, for example). Failing that, *FreeText* will tell the learner what type of error has been made (e.g. an agreement error in number). As a last resort, the first two levels of feedback will be combined so that the learner can simultaneously see the location of the error and be informed of the type of error that has been made.

In their ILTS for learning the passive voice in Japanese, Yang & Akahori (1997) do not give details about the content of their feedback or how many levels of it there are, but they do mention that it becomes more detailed with every correction. In addition to the feedback, the ILTS refers the learner to hyperlinks to reference material relevant to the error, instead of providing the correct answer.

Some ILTSs that incorporate several levels of feedback have the additional capability of individualizing the messages to the learner's level of competence. To do this, a score is kept of the learner's performance on each error type and stored in the learner model. The score is automatically and continuously updated as the error is committed or avoided and the level of feedback is modulated accordingly. For example, the *German Tutor* (Heift 2001) provides three levels of feedback for each error type which range from identifying and explaining the error (beginner), to stating the location of the error (intermediate), to advising that there is an error in the sentence (advanced). Depending on how many errors the learners make as they work through the tasks set by the *German Tutor*, they could be at the beginner level for one area of grammar (e.g. case assignment) but at the

---

[3]*FreeText* was developed as part of a large-scale project partially funded by the European Commission. Each of the four contributors (Université de Genève, University of Manchester Institute of Science and Technology (UMIST), Université Catholique de Louvain and Softissimo SA) brought their own expertise to the project, resulting in a number of publications, in addition to L'haire (2004) cited above, to do with the NLP tools, the materials and the learner corpus used to create the error taxonomy exploited by the software. For a list of these publications, see the project website at http://www.latl.unige.ch/freetext/en/publications.html

advanced level for another (e.g. subject-verb agreement). Furthermore, the levels will be constantly adjusted so that the learner who is at the beginner level for case assignment will be moved to the intermediate level as they make fewer case assignment errors.

Criswell et al. (1992) describe a prototype ILTS which also uses feedback individualized at three levels: maximal, medium and minimal assistance users. They do not offer much detail regarding the content of the messages or the error types that their ILTS can handle. However, they do mention that a maximal assistance learner will receive feedback that is not only more elaborate, but also more frequent and that the learner model will reclassify the learners' level according to their progress.

On a given error, *CASTLE*[4] (Krüger & Hamilton 1997, Murphy & McTear 1997), an ILTS for learners of English, offers two levels of feedback before providing the correct answer and generating remedial exercises. In addition to tracking the number of times an error is made, the number of times that the learner refers to the built-in help functions is also considered in determining at what point the level of feedback should change.

XTRA-TE (Chen & Kurtz 1989) was developed out of a Chinese to English machine translation system. For each error type, learner knowledge is characterized as belonging to one of four levels. At the first level, the learner is asked to self-correct without being told where the error is in the sentence. At the second level, *XTRA-TE* provides a hint about the error and asks the learner to try again. Failing that, the error is explicitly identified and the correct form is given. At this point, if the learner is still not able to produce the correct form, *XTRA-TE* explains the error and provides remedial exercises on the relevant area of grammar.

In the two approaches to the provision of feedback and indeed among all, except Zock et al. (1989), of the ILTSs discussed thus far, the feedback has been explicit. Another strategy has been to implicitly indicate that there is an error either by asking the learners questions about their errors or by stating the rule that has been violated.

---

[4]*CASTLE* was part of the RECALL (Repairing Errors in Computer Assisted Language Learning) project funded by the European Union's Telematics Applications of Common Interest - Language Engineering LE1-1615. Krüger & Hamilton (1997) refer to the ILTS as RECALL, rather than as *CASTLE.*

Miniprof (Labrie & Singh 1991), for example, is designed to prompt the learner to self-correct errors of negation, pronominalization and conjugation in French. For instance, they are asked to answer the question *"Est-ce que Mike étudie beaucoup?"* (*Does Mike study a lot?*) by replacing the proper noun *"Mike"* with a pronoun. If they do not do so correctly, *Miniprof* will reply by saying that the learners need to substitute a pronoun for "Mike" and then it will ask: *"Which pronoun should you use?"* (Labrie & Singh 1991, p. 11). If the learners either do not know which pronoun to use or use the wrong pronoun, *Miniprof* tells them what the appropriate pronoun is and presents the correct sentence in its entirety.

Rather than stating directly that an error has been made, the PAS (Prompting Answer Strategy) feedback used by Ferreira-Cabrera (2003) in her ILTS for the Spanish subjunctive gives the rule that has been broken in the learner input. For example, if the learners incorrectly use the indicative mood following a main clause expressing advice, they will receive a feedback message such as this: *"After verbs which advise or suggest (e.g. Aconsejaría), you have to use the subjunctive mood."* (Ferreira-Cabrera 2003, p. 161).

ICALL Summary and Evaluation

Previous work in ICALL has produced ILTSs whose focus has been on innovations in NLP, pedagogy or feedback. Of those in which feedback has been a central issue, three main approaches have been identified. Firstly, to encourage the learner to self-correct, researchers have proposed and implemented the idea of using several levels of feedback, each with an increasing degree of specificity about the error. Secondly, to tailor the feedback to the appropriate level for each learner, ILTSs have been devised which dynamically modulate the feedback level by error type. Thirdly, some ILTSs implement an implicit approach whereby the learners are not told explicitly that there is an error, but rather are asked a question or provided with a violated rule which is intended to lead them to realize their mistake.

Of the 18 ILTSs reviewed above, none has undergone any quantitative testing that would confirm or refute the relative effectiveness of ILTSs generally, nor

of intelligent feedback in particular. In fact, to our knowledge, work by Nagata (1993, 1995, 1996, 1997a,b, 1998), Nagata & Swisher (1995) and Ferreira-Cabrera (2003), which will be discussed in detail in Section 2.4, is the only existing research that addresses questions of effectiveness. More quantitative work is clearly needed in order to establish what, if any, types of intelligent feedback have an effect on the language learning process.

While the approach to grammar teaching in some ILTSs has drawn on language learning and teaching theories, the same cannot be said of feedback. With the exception of Heift (2001), none of the feedback in the ILTSs discussed above is theoretically motivated. Some researchers (e.g. Levy (2000)) take the position that CALL should build on itself as an independent field of inquiry, rather than borrowing from other disciplines. While we accept the validity of such an approach, the question arises as to why it would not be advantageous to make use of existing knowledge about second language acquisition. One response by Chapelle (2005) has been to suggest that the theoretical contribution from related fields, such as instructed second language acquisition, is not extensive enough to be able to inform the design of software for language learning. This is true to the extent that a theory of second language acquisition is, at the moment, out of reach (Gregg 2003). However, since the emergence of modern second language acquisition as a field in its own right in the late 1960s, there have been significant advances concerning the acquisition process (Long & Doughty 2003). As the review of the second language acquisition literature on feedback in the next section will show, there is certainly a substantial enough base of knowledge on which to draw. Furthermore, second language acquisition does not need to be regarded as an authoritative source of knowledge in order for it to be useful in (I)CALL. On the contrary, rather than assuming a unidirectional flow of knowledge, we advocate the view that questions which are of interest in second language acquisition and (I)CALL should be investigated with a view to advancing knowledge in both fields.

2.3 Feedback in the Second Language Acquisition Literature

In their synthesis and quantitative meta-analysis of research on the effectiveness of second language instruction, Norris & Ortega (2000) identified six over-

arching research questions that this literature addressed between 1980 and 1998. One of them was whether or not negative feedback is beneficial for second language development and, if it is, what types of feedback may be most effective. Accordingly, feedback was tested in the relevant studies independently of instructional treatment. This work, as it relates to errors in morphology and syntax, will be reviewed in Section 2.3.2 below.

2.3.1 Descriptive Interaction and Classroom Research on Feedback

It was not until the 1980s that the definition of negative evidence in the first and second language acquisition literatures was expanded from a simple correction to include a number of phenomena, such as confirmation checks, clarification requests, failures to understand, silence and expanded and corrected repetitions (Schachter 1991)[5]. The expansion of what constitutes negative evidence in second language acquisition initially came as a result of observational studies of naturalistic conversations between native speakers and second language learners (see, for instance, Chun et al. 1982, Crookes & Rulon 1988, Day et al. 1984, Gaskill 1980).This work continued and subsequently grew into an area of research that explores the role of negotiated interaction in non-native speaker dyads as well as between native speakers and non-native speakers (Gass & Varonis 1985, 1989, Pica 1988, Pica et al. 1986, 1987, Varonis & Gass 1985) and has culminated in the Interaction Hypothesis (see Section 2.3.2 below). Several studies have also described the feedback provided to learners by teachers in classroom settings (e.g. Allwright 1975, Courchene 1980, Holley & King 1971, Long 1988, Swain & Carroll 1987) and, in some cases, have resulted in elaborate feedback taxonomies (Chaudron 1977, Lyster & Ranta 1997).

Some of this descriptive research has led to interesting hypotheses concerning what feedback might be most beneficial to learners. For instance, from analysis of their classroom observational data, Lyster & Ranta (1997) found that, while on average recasts accounted for 55% of the feedback provided to learners by teachers, they led to substantially less learner-generated repair compared to

---

[5]We acknowledge here that there is a parallel literature on feedback in first language acquisition. See Long (1996, p. 430-437) for a review.

elicitation, metalinguistic feedback, clarification requests and repetition. Consequently, the authors suggested that perhaps recasts are not a particularly effective type of feedback. This is clearly a hypothesis, rather than a conclusion, as Lyster and Ranta did not measure learning gains resulting from recasts, nor any of the other types of feedback that they describe in their study. Commenting on interaction research generally, Gass (2003) acknowledges the difficulty in proving that the negotiations that occur during interaction result in learning. Therefore, descriptive interaction and classroom research will not be reviewed here and, instead, we will turn our attention to studies whose purpose has been to test the effectiveness of feedback.

2.3.2 The Interaction Hypothesis and Focus on Form

A large amount of the research that has sought to test the effectiveness of feedback is motivated by the Interaction Hypothesis proposed, and in later work updated, by (Long 1981, 1983, 1996) and its pedagogically-oriented counterpart, *Focus on Form* (Long 1991). The basic tenet of both constructs is that drawing attention to linguistic form will be most useful in second language development when the focus of interaction is on meaning. It follows that providing feedback, particularly in the form of implicit negative evidence, may be a means of achieving this goal. Consequently, a number of studies have set about testing the effectiveness of feedback according to the Interaction Hypothesis/Focus on Form framework. At a theoretical level, this framework has been subject to criticism by Carroll (2001), Gass & Varonis (1994) and Loschky (1994) for its lack of explanatory power. Specifically, the Interaction Hypothesis does not explain what needs to happen in the mind of the learner for the linguistic form to be acquired, nor how focusing on form in an interactive context is supposed to help in the acquisition process. These questions, as they relate to our experiment, will be taken up in detail in Section 5.2. Despite the theoretical vagueness of the Interaction Hypothesis, the experimental portion of the literature represents a sizable body of work on feedback. As such, it will be reviewed below by feedback type and combinations thereof.

Motivated by a need to test both the feasibility and the effectiveness of focus on form in instructed Second Language Acquisition, Doughty & Varela (1998) investigated the effect of corrective recasting on the acquisition of the simple past and conditional tenses by 34 predominantly Spanish-speaking students aged between 11 and 14 in the setting of two intermediate English as a Second Language (ESL) science classes. They used a pretest-posttest design, in which one class received feedback throughout the treatment phase and the other did not. A delayed post-test to measure any long term effects of the feedback was also carried out two months after the treatment.

Corrective recasting, involved two steps: firstly, the teacher would repeat the learner's utterance, possibly using rising intonation and stressing the non-target-like verb form, in order to signal the mistake. Secondly, if the learner made no attempt to self-correct, the teacher would recast the utterance, stressing the target-like form of the problematic verb and finishing with falling intonation. For clarity, the example cited in Doughty & Varela (1998, p. 124) showing both steps of the corrective recast is reproduced below. In this case, the learner has produced an utterance in the present tense when the context required the past tense:

(2.1)   Corrective Recasting
  José: I think that the worm will go under the soil.
  Teacher: I *think* that the worm *will* go under the soil?
  José: (no response)
  Teacher: I *thought* that the worm *would* go under the soil.
  José: I *thought* that the worm *would* go under the soil.

Six different science experiments served as the content for the experimental materials. Lab reports on experiments 1, 5 and 6 were used as pre, immediate post and delayed posttest while the rest were used as part of the treatment phase, in addition to a number of other tasks. These additional tasks included a class discussion in which the students responded to questions about the science experiments posed by the teacher that were designed to elicit the target tenses and presentations of the procedure followed in their science experiment to the class

and to each other in small groups. In the lab reports, the students answered four questions about the experiment, three in the past tense and one in the conditional, in writing and were interviewed by the teacher using the same questions, as well as an additional one in the conditional, all of which the students were to answer spontaneously while being audiotaped. These data were used to measure both written and oral improvement in producing the simple past and conditional tenses.

The pre and post test data were coded as target-like, non-target-like (e.g. using the present tense where the past was appropriate) and interlanguage (i.e. identifiable marking of the target-like function using a non-target-like form). Substantial gains by the *focus on form* group were made on both the written and oral measures on the immediate posttest. In contrast, the control group made no improvement with the exception of a small, but significant, increase in the emergence of interlanguage past time reference on the written test. The authors attributed this improvement to independent progress on the part of the learners.

The results on the oral measure were maintained two months subsequent to the experiment by the *focus on form* group. As for the written test, the results for target-like marking did not endure and, while the interlanguage and non-target-like results did, they were not as strong as for the oral test. The authors proposed two possible, yet contrasting, explanations for the results of the written tests. Focusing on the possible weakness of the feedback, they suggested that, as corrective recasting was conceived as an oral form of feedback, perhaps it was less appropriate for the correction of writing. Yet, advocating the usefulness of corrective recasting, they also claim that the consistency of the feedback in both writing and speaking may have had a positive overall effect. The explanation for the results of the written test aside, due to the control group's lack of progress between the immediate and delayed post-tests, the authors were able to conclude that overall corrective recasting helped learners towards target-like production of past time reference.

Ishida (2004) also investigated the effect of recasts as a *focus on form* technique, however hers did not contain the corrective move used by Doughty & Varela (1998) of drawing attention to the learner error by repeating it with rising intonation before delivering the recast. The Japanese aspectual form *-te i-(ru)* was

chosen as the target structure, as a further motivation for the study was to test the predictions of the Aspect Hypothesis. Four learners of Japanese participated in the study, three of whom were native speakers of English and one of whom was a Chinese-English bilingual[6]. The procedure consisted of eight one-on-one conversation sessions between each subject and the investigator on the topics of daily activities, weekly schedules, a picture description, weekend talk and reflection on the session. Following a time-series design, two of these sessions served as pre tests which were a benchmark against which four treatment sessions and two post test sessions were compared.

The results showed a significant increase in scores from pre to post test sessions which was maintained in the delayed post test seven weeks after the treatment. There was also a significant positive correlation between number of recasts and target-like usage. However, it was also found that there was a high increase in accuracy in the resultative use of *-te i-(ru)* with relatively few recasts, compared to the number supplied for progressive and habitual uses. The author accounts for this by suggesting that the effectiveness of recasts is constrained by factors such as learner readiness, prior instruction and the complexity of form-meaning mapping with respect to the target structure. While this may be true, since there was no control group and the experimental participants were attending Japanese classes at the time the experiment was being conducted, any improvement cannot necessarily be attributed to the feedback. Moreover, the author makes the mistake of using the correlation results to claim that the feedback *causes* improvement in target-like usage. This claim cannot be substantiated because of the third-variable problem; that is, that there may be other variables, aside from the feedback, which contributed to the improvement in target-like usage (for further explanation of the third variable problem, see Field 2005, p. 128).

Mackey & Philp (1998) were also interested in the effect of recasts on second language development, but in the context of native speaker-non native speaker task-based interaction, rather than *focus on form* instruction. In light of the tendency for previous studies to look only at immediate effects, the two goals of their experiment were, firstly, to test recasts to see whether an effect one week and four weeks following treatment was evident and, secondly, to explore how

---

[6]The author acknowledges the limitations of the study with respect to the small sample size as well as the advantage that the Chinese-English bilingual had over his anglophone counterparts.

learners responded to recasts. The developmental readiness of the subjects to learn the target structure of question forms, according to Pienemann & Johnston (1987) and Pienemann et al. (1988)[7], was also taken into account.

Thirty-five ESL learners at a beginner and lower intermediate level participated in the experiment. They engaged in picture drawing, story completion and story sequencing tasks[8] which were designed to necessitate interaction using question forms for 15-25 minutes once a day for a week. Pre, immediate post and two delayed post tests were conducted using a picture difference task[9]. The authors found that learners who were developmentally ready and received recasts were able to produce more question structures at higher developmental levels than those who did not receive recasts. The results also showed no relationship between uptake (i.e. learners' immediate responses to recasts) and short-term development.

Han (2002) explored the question of whether recasts are effective in improving linguistic forms over which learners are developing processing control, as opposed to forms for which learners are at the stage of forming mental representations. With this in mind, her study tested the effects of recasts on learners of English who showed a lack of consistency in producing the present and past tenses. The experiment used a pre-post test design and involved eight learners, four of which received recasts in instances where the past tense was not used consistently. For the pre and post tests as well as for the experimental treatment, the subjects were asked to narrate the stories depicted in a series of cartoon strips, first in writing and then orally. Recasts were supplied exclusively on oral production and the treatment took place twice a week for four weeks.

Comparisons of proportional mean scores showed that the recast group produced tenses more consistently than the control group and that the improvement was maintained over the period of a month between the immediate post test and delayed post test. The results are rather misleading, however, in that the recast

---

[7]This work identifies six stages through which learners progress in acquiring question formation in English as a Second Language. They are part of a larger line of research known as Processability Theory (Pienemann 1998, 2005) that predicts which language forms learners can process at a given developmental stage.

[8]There are no details of what exactly these entailed.

[9]For this task, both the NS and NNS are given two pictures that are similar but include ten differences. The NNS's job is to find the differences between the two pictures by asking the NS questions about his or her picture.

group is evaluated on the basis of consistency in *past tense* production whereas the control group is evaluated according to their *present tense* production. From the beginning of the experiment, the control group shows a clear preference for the present tense in both oral and written narrations. While this trend continues for the duration of the experiment in their oral production, the learners start to use the present tense less and the past tense more in their writing. As a result of the type of analysis adopted for this study, the control group is penalized for this shift when in fact it may well be that they were actually just developing their competence in past tense production. While the recast group does appear to have made substantial improvement, without a reliable control measure, their progress could be attributed to instruction that they were receiving in class during the period of the experiment and not necessarily to the feedback.

Clarification Requests

Nobuyoshi & Ellis (1993) conducted a study which tested whether clarification requests would contribute to the acquisition of the past tense by six Japanese-speaking learners of English. The tasks consisted of describing what happened the previous weekend and the previous day at the office according to a series of pictures. There were two sessions, between which there was a one-week interval. During the first session, the experimental group received clarification requests following learner utterances in which the verbs were either not in the past tense or were incorrectly formed. During the second session, the experimental group received clarification requests exclusively in response to utterances that were genuinely unintelligible. These "genuine" clarification requests were given to the control group for both sessions of the experiment.

Comparisons of scores in sessions one and two showed that learners 1 and 2 of the experimental group improved as a result of the feedback, going from an accuracy rate of 31% to 89% and 45% to 62% respectively. Learner 3, however, made no improvement, nor did any of the members of the control group. The authors interpreted the results as providing support for the idea that clarification requests, as a means of "pushing" learners to improving the accuracy of their language production, contribute to improving performance both immediately and over time for some, but not necessarily all, learners. However, without systematic pre and post testing and a more substantial sample size, the results

could also be attributed to differential levels of competence across learners or individual variation.

Positive Evidence versus Negative Evidence

Some of the studies which have been motivated by the Interaction Hypothesis have also considered the effect of positive evidence in contrast to negative evidence.

Long et al. (1998) reports on two studies, one on Japanese and the other on Spanish[10], which looked at the effects of models and recasts on previously unknown second language target structures. The Japanese study used adjective ordering and fronted locative phrases as target structures. For the adjective ordering, both the word order colour-size-NP and the addition of the gerundive form, *-kute*, to the first of two juxtaposed adjectives were targeted. The following example, taken from Long et al. (1998, p. 360) is reproduced below to illustrate:

(2.2)   Adjective Ordering Target Structure
          aka-kute ookii hako
          red-gerund large box
          a large red box

Thus, in their data analysis, the researchers took into account whether the learners produced adjectives of colour followed by adjectives of size as well as whether they added *-kute* to the first adjective in the sequence. The target locative construction was the one preferred when there is no established topic in the discourse. To express this, the locative phrase occurs before the subject, rather than between the topic and verb, as it does when a topic either is being, or has been, introduced.

To test the target structures, each participant and the researcher participated in a communication game. For both structures, the learner produced an utterance describing a visual stimulus which evoked the use of the target structure. For the adjective ordering structure, the stimulus was one of eight pieces of paper, chosen by the learner, which were large or small and either blue, red, black or white.

---

[10]These studies have been published in greater detail in separate articles. See Inagaki & Long (1999) and Ortega & Long (1997).

For the locative construction, the learners described the position of their choosing of four dolls on a drawing of a room with two rows of seats. The researcher had to select or order the visual stimulus, as it was described by the learner, and both participants, separated by a screen, held up their stimuli to confirm that they matched. In the recast condition, the researcher recasted the learner's utterance, whether it was target-like or not, before comparing stimuli. In the model condition, the learners were given the utterances through headphones which they then repeated to the researcher. The control group spent the treatment time practicing kanji characters that were unrelated to the target structures. A picture description task was used to measure improvement in production.

The results showed that there was no significant difference between the experimental groups and the control, nor between the two experimental groups for either target structure. However, results with respect to the gerundive form *-kute* in the adjective ordering condition showed that both of the experimental groups outperformed the control group, but did not outperform each other.

The Spanish study looked at object topicalization and adverb placement. Object topicalization occurs in Spanish when the object is an established referent which the speaker wants to emphasize. In the syntax, it is manifested by placing the object in pre-verbal position and adding a pronominal copy between the object and the verb. Compare below a sentence in canonical word order in example 2.3 to the same sentence with a topicalized object in example 2.4 (Long et al. 1998, p. 364):

(2.3)   Canonical word order
        Pepe tiene la guitarra
        Pepe has the guitar

(2.4)   Object topicalization
        La guitarra la tiene Pepe
        The guitar it has Pepe

For adverb placement, the particular structure tested was that which occurs in Spanish but not in English, namely SVAO (Pepe toca todos los días la guitarra but *Pepe plays every day the guitar).

The treatment sessions were basically the same as in the Japanese study with some minor modifications to take into account the different target structures. A picture description task and a grammaticality judgment task were used as pre and post test measures of production and comprehension. No feedback effects were found with respect to object topicalization. However, recasts and models both had an effect on the learners' proficiency in adverb placement, particularly for the recast group who outperformed the model group.

In addition to models and recasts, Ayoun (2001) used traditional grammar instruction as an experimental condition which provided a combination of explicit positive evidence and negative feedback. She predicted that recasts would be more effective than models and traditional grammar instruction in helping learners towards target-like production of the aspectual distinction between the *passé composé* and the *imparfait* in French. The experiment took place over a period of five weeks. During the first and last weeks, pre and post tests, consisting of a composition task, were administered. The pre test was used to establish a uniform level of competence across learners along with a grammaticality judgment test for which the learners had to score sentences as grammatical on a scale of one to five, as well as to provide corrections for sentences they considered to be incorrect.

The traditional grammar instruction group was presented with information about aspect, such as how to conjugate the tenses, how to use them, what the aspectual distinctions were and how to choose between them, during each of the three treatment sessions. This was followed by practice exercises in the form of contextualized passages that were as similar as possible to what the other groups received which gave the learners a chance to practice what they had just learned. Following the practice, they were given an answer key to check their answers. The recast group did a gap-fill exercise on sentences which corresponded to pictures that created a story. After the response was entered, the correct answer was shown for three seconds, whether the response was correct or not[11]. The modelling group were given a model sentence of what was depicted in the picture for three seconds and then asked a related question about the story. As the experiment was computerized, all feedback was delivered in writing, rather than orally.

---

[11]The author concedes that for this reason recasts did not exclusively constitute negative evidence. Whether recasts qualify exclusively as negative evidence at all is addressed by Leeman (2003) further on in this section.

After each treatment session, all groups did an additional to task to monitor their progress. These included a translation from English into French, a grammaticality judgment task and a preference task and were meant to be different from each other so as to avoid task effects.

As predicted, all groups improved significantly from pre to post test, although there was no control group to confirm that this was due to the feedback used in the study and not an effect of classroom instruction throughout the period of the experiment. The recast group performed significantly better than the traditional grammar instruction group, but not the modelling group. The results also showed improvement in production of the *passé composé* across all groups. This was not the case, however, for production of the *imparfait*, as the recast group performed better than the modelling and traditional grammar instruction groups.

In addition to exploring interaction effects and the type and frequency of feedback that native-speakers provide to learners, Iwashita (2003) analyzed the relationship between native-speaker feedback and learner improvement from pre to immediate post test on the development of word order and particle use in fronted locative phrases as well as *te*-form verbs[12] in Japanese. She analyzed four types of naturally occurring native-speaker feedback: recasts and negotiation moves[13], which were instances of negative evidence, and simple models and completion models, which were occurrences of positive evidence in her data. She found that recasts significantly contributed to improvement on the *te*-form verb, but not on either word order or particle use in sentences using fronted locative phrases. There was also an effect for simple models on all three target structures, but only for learners who had an above-average score on the pre test. It should be noted, however, that feedback type and pre test score only accounted for 23-38% of the variance in immediate post test scores.

In an effort to explain the contribution of recasts to language development, some researchers have proposed that the juxtaposition of the learner utterance and the recast, which differ only slightly, will enhance the salience of the different element(s) and promote rejection of the non-target-like form(s) (see Farrar (1990),

---

[12]The author characterizes *te* is an inflectional morpheme used on some, but not all, verbs to indicate commands and requests.

[13]Negotiation moves are defined as "an interactional move intended to clarify the meaning of a NNS's non-target-like or incomplete utterance" (Iwashita 2003, p. 13).

Nelson (1987), Saxton (1997) for first language acquisition and Long (1996), Long et al. (1998) for first and second language acquisition). Leeman (2003) based her experiment on the idea that it is the highlighting of the target form(s) by recasts, in other words positive evidence with enhanced salience, that may be the beneficial element of recasts. Accordingly, she sought to disentangle the contribution of recasts to second language development as being attributable to negative evidence, unenhanced positive evidence or positive evidence with enhanced salience using adjective-noun agreement in Spanish as a target structure. The study incorporated four conditions: (1) negative evidence (2) positive evidence with enhanced salience (3) recasts, which are a combination of 1 and 2, and 4) unenhanced positive evidence, which served as the control.

Seventy-four anglophone learners taking first-year undergraduate courses in Spanish volunteered to participate. For the pre, immediate post and delayed post test one week later, they were asked to perform picture-difference tasks for which they had to identify as many differences as they could between two similar photographs to a maximum of 32 items per test. The photographs differed in the colour, number and placement of various objects, rendering the production of sentences involving noun-adjective agreement a necessary part of the task.

The treatment tasks consisted of an object placement task and a catalogue-shopping task which were carried out in two phases. In the first phase, the learners had to give directions to the researcher, and in the second they had to follow directions from the researcher. An example of each feedback type is given below in response to a gender agreement error, where the adjective modifying the feminine noun, *taza*, has been marked as the masculine, *rojo*, when it should be the feminine, *roja*. In the event of such an error, the recast group received a target-like version of the offending noun phrase, as follows (Leeman 2003, p. 48):

(2.5) Recast group
  NNS: *En la mesa hay una taza *rojo.*
      "On the table there's a *red cup."
  R: *Um hmm, una taza roja. ¿Qué más?*
      "Um hmm, a red cup. What else?"

The negative evidence group received a feedback message that pointed out the error (Leeman 2003, p. 49):

(2.6)   Negative evidence group
          NNS: *En la mesa hay una taza *rojo.*
                  "On the table there's a *red cup."
          R: *Um hmm, pero tú dijiste 'una taza *rojo.' ¿Qué más?*
                  "Um hmm, but you said 'a *red cup.' What else?"

The enhanced salience and control groups received no feedback on adjective-noun agreement in this phase:

(2.7)   Enhanced salience and control groups (phase 1)
          NNS: *En la mesa hay una taza *rojo.*
                  "On the table there's a *red cup."
          R: *Um hmm. ¿Qué más?*
                  "Um hmm. What else?"

However, in the second phase, the enhanced salience group received input in the form of the researcher's directions which used stress and intonation to highlight the inflectional morpheme on the adjective (Leeman 2003, p. 49):

(2.8)   Enhanced salience group (phase 2)
          R: *La manzana rojA está en la mesa.*
                  "The red apple is on the table."

Meanwhile, the rest of the groups received the same input without any enhancement:

(2.9)   Other groups (phase 2)
          R: *La manzana roja está en la mesa.*
                  "The red apple is on the table."

The results revealed that only the recast and positive evidence with enhanced salience groups showed a significant improvement in performance as compared to the control group. This was true for both gender and number agreement on

both post tests with the exception of gender agreement on the second post test which was not significant for the recast group. The negative evidence group did not perform significantly better than the controls. The author concluded that negative evidence is not as crucial as positive evidence with enhanced salience in having a beneficial effect on second language development.

In keeping with the Interaction Hypothesis, feedback that is given during inter-action should successfully draw the learner's attention to linguistic form with minimal disruption to interaction. Not surprisingly, the three types of feedback that have been the object of study of all the research outlined above - clarifica-tion requests, models and, in almost every experiment, recasts - fit this profile. We turn our attention now to the comparatively few studies that have explored other types of feedback, such as metalinguistic explanations, explicit utterance rejection and indirect metalinguistic feedback.

2.3.3   The Psycholinguistic Studies

Kubota (1997) investigated the effect of feedback on the acquisition of phrasal verbs among Japanese learners of English. Feedback consisting of a) positive ev-idence, b) positive evidence and negative evidence with explicit metalinguistic information and c) explicit response rejection and modelling were tested on six sentences for a total of 96 learners. In the positive evidence condition, gram-matical versions of the sentences using phrasal verbs were written on the black-board and read by the teacher three times each. In the positive and negative ev-idence with explicit metalinguistic information condition, the grammatical and ungrammatical versions of the same six sentences were presented to the learners. They were then given metalinguistic rules for phrasal verb usage. In the explicit response rejection and modelling condition, learners worked individually with the researcher. They had to state whether the sentences were grammatical or not and were told when their judgments were wrong and given a model.

The pre and post tests consisted of a grammaticality judgment test and a trans-lation test. There were post tests immediately after the treatment and at sub-sequent intervals of one month and one year. On the grammaticality judgment measure, the results showed that all of the experimental groups outperformed the control group and, furthermore, that the metalinguistic information group

outperformed the other feedback conditions both immediately following and one month after the treatment. On the translation measure, only the metalinguistic information group consistently outperformed both the control and the other feedback conditions immediately following and one month after treatment. The positive evidence group showed significant immediate improvement but not after a month. The explicit response rejection and modelling group experienced no improvement at all. While overall the results show a clear advantage for the metalinguistic information group, identical tests were used to measure improvement and a subset of the questions on these tests were used as target items in the experimental treatment. Furthermore, rather than receiving treatment without feedback, the control group received no treatment at all. Therefore, the possibility that the results reflect successful memorization of the individual target sentences, rather than a learning gain due to the provision of feedback, cannot be excluded.

Carroll et al. (1992) investigated the effectiveness of explicit correction on the learning of two nominal suffixes in French, *-age* and *-ment*. Subjects included 79 native speakers of English split into intermediate and advanced groups. After being trained on the relationship between two verb types and their related nouns, the learners read a series of sentences in French written on cards with English translations. Each sentence contained a graphically highlighted example of one of the two verb types. The learners then were to state the appropriate noun corresponding to the verb in the sentence.

In the first stage of the experiment, the treatment group were given explicit correction (i.e. explicit rejection followed by a model of the target noun) while the comparison group was not. In the second stage, the subjects were given completely new stimuli and no feedback or correction was given to either group. The results of the first stage showed that the treatment group performed significantly better than the control group. However, this was not the case for the second stage. Carroll et al. (1992) concluded that error correction had an effect but that it did not help learners to construct morphological generalizations. Rather, correction aided them in memorizing lexical items.

Using the same methodology, Carroll & Swain (1993) looked into whether 100 hispanophone learners of English could induce the phonological and semantic

constraints on double object verbs[14] (i.e. V+NP+PP (e.g. *The woman gave a present to the boy*) or V+NP+NP (e.g. *The woman gave the boy a present*)). The constraints are that the verb stem must consist of one syllable or of two syllables with stress on the first and that the first object in the construction must be a recipient and therefore [+animate]. As part of the instructions for the experiment, the learners were told that they would be given a sentence and their task was to find a different way of saying the same thing. They were given examples of a transformation and told that the original sentence and the transformation were equivalent in meaning.

Errors made by the experimental groups were corrected using one of four types of feedback: *explicit hypothesis rejection*, which involved being told they were wrong and being given a phonological or semantic rule for double object verbs, *explicit utterance rejection*, meaning that they were told that their utterance was incorrect, *modelling plus implicit negative feedback* (or recast) and *indirect metalinguistic* feedback, in other words being asked if they were sure of their response. All groups were told that they would receive feedback when they made a mistake. The control group received no feedback.

The results for both the feedback and the guessing sessions showed that all of the experimental groups performed significantly better than the control group and that the group which had received *explicit hypothesis rejection* feedback outperformed the other groups overall. The authors concluded that in addition to helping the subjects learn the stimuli, the feedback, and in particular the *explicit hypothesis rejection* feedback helped them to identify the correct generalization. Kubota (1994) replicated this study among Japanese learners of English and found that only the *modelling plus implicit negative feedback* group learned the experimental stimuli and that none of the feedback groups were able to come to any generalizations about how dative alternation worked.

---

[14]This phenomenon is also known as dative alternation and has been the subject of studies in the second language acquisition literature by Hawkins (1987), Mazurkewich (1984) and White (1987) addressing questions of markedness and transfer. However, it is the focus of this study by Carroll & Swain (1993), as well as others (Bley-Vroman & Yoshinaga 1992, Inagaki 1997), following work in the first language acquisition literature which used dative alternation as a means to test whether or not overgeneralization can be avoided without the provision of negative evidence (we refer the reader back to Section 1.1.3 for further discussion). Dative alternation has also been a topic of investigation in the language attrition literature in work by Seliger (1989, 1991).

A third study looked at the category conversion of verbs into nouns in English (Carroll 2001). It followed the same design and methodology of the previous two experiments that Carroll was involved in and incorporated the four feedback types as well as the same subjects used by Carroll & Swain (1993). The examples below, taken from Carroll (2001, p. 322), illustrate the contrast between the use of the word *help* as a verb and a noun and show that in English there is no morphological marking to indicate which category the lexical item belongs to:

(2.10)   Can you *help* Margarita with her homework?
          I think she needs some *help*.
          Many people *help* me at odd jobs.
          I am grateful for this *helping* out.

The objective was for the subjects to learn how to derive nouns both with and without a derivational *-ing* suffix from an uninflected verb stem.

The results were similar to Carroll & Swain (1993) in that all of the experimental groups were able to learn the stimuli on the basis of the feedback they received. However, for the novel stimuli, only the *explicit hypothesis rejection* and the *indirect metalinguistic* feedback groups exhibited significant learning gains as compared to the control group. The conclusion was that only these two feedback types helped learners towards forming a linguistic generalization about how to convert verbs to nouns.

2.3.4   *Feedback in Second Language Acquisition Summary and Evaluation*

Judging by the studies reviewed here, recasts have become the feedback type of choice among proponents of the Interaction Hypothesis. Research on recasts have been shown to be effective in promoting second language development of past simple and conditional tenses (Doughty & Varela 1998) and question forms in English, subject to developmental readiness (Mackey & Philp 1998). Compared to other types of feedback, the results for recasts have been mixed. Long et al. (1998) found little effect of recasts or models on the locative and adjective order structures in Japanese, no effect on objective topicalization but significant effects for models over controls and recasts over models for adverb placement in

Spanish. Ayoun (2001) found overall that models and recasts were more effective than traditional grammar instruction but not more effective than each other for developing production of past tense aspect in French. However, on the individual tenses, recasts were more beneficial than any other feedback type for producing the imperfective, while all three feedback types had an effect on the perfective. Iwashita (2003) reported stronger results for models than for recasts on the *te*-form verb and locative construction in Japanese. Results from Leeman (2003) favoured recasts and positive evidence with enhanced salience over negative evidence for adjective-noun agreement in Spanish.

In comparison, research on more intrusive feedback types has been limited to work by only a handful of authors. Results thus far have also been mixed but the effectiveness of explicit deductive feedback (i.e. that which explains the error using a metalinguistic rule) appears to have met with some success in helping learners to form linguistic generalizations. As the target structures in question (i.e. phrasal verbs, double objects and category conversion in English) function according to two or three basic rules, explicit deductive feedback is feasible. However, there are other areas of grammar where what is expressed in the morphology and syntax is governed by complex semantic and pragmatic constraints. Aspect, the target structure of the experiment conducted for this thesis, is one such area. Reducing its constraints down to a few rules is not straightforward. Moreover, attempts to do so, in French grammar books at least, typically result in inaccurate and oversimplified rules (Andrews 1992). Thus, perhaps explicit *inductive* feedback, whereby in the face of an error, the learner is given the exact interpretation for a specific sentence, rather than being presented with a rule that embodies a particular type of interpretation, would be a type of feedback worth investigation. Further detail on how this would work for aspect in French will be discussed in Section 5.4.1.

2.4   Evaluative Studies of ICALL Feedback

The sum total of work that has sought to evaluate the effectiveness of ICALL feedback has been conducted by Nagata (1993, 1995, 1996, 1997a,b) and Ferreira-Cabrera (2003). Their research will be reviewed in detail in this section.

Motivated by a need to justify intelligent computer-aided language instruction (ICALI)[15] and to determine the type of feedback that it should provide, Nagata (1993) conducted an experiment which compared the effect of traditional versus intelligent CALI feedback using *Nihongo-CALI*, a purpose-built programme written by the author[16]. The target structure was the production of passive sentences in Japanese and was tested on 34 subjects, all of whom were taking a second-year Japanese language course at the University of Pittsburgh at the time of the experiment.

The majority of the experiment took place in the space of a week and consisted of six one-hour sessions, the first and last in the classroom and the remaining four using the software. In the first session, the subjects were given brief in-class explanations of the passive voice and wrote a test on basic Japanese grammar. Rather than serving as a pre test, the purpose of this assessment was to evaluate their grammatical knowledge so that they could be paired and randomly assigned to one of the two feedback groups. During each of the following four computer sessions, the learners read grammar notes and then practiced one of the three passive constructions (i.e. the direct, indirect and honorific) using *Nihongo-CALI*. The fourth session served as a review of all three constructions. In the final session of the week, the subjects were post tested in class using a test of the same format and content as the experimental exercises. Four questions on the passive voice which appeared on the end of semester exam served as a delayed post test three weeks following the experiment[17].

For the experimental task, the subjects were presented with a situation involving them and an interlocutor. This was followed by a written statement or question made by the interlocutor to which the learner had to respond in the passive voice, as the following example from Nagata (1993, p.335) shows:

---

[15]Computer-Assisted Language Learning has become the generic name for the field. However, ICALI, among others terms, was used in the past and hailed from researchers who were more focused on teaching than learning. For discussion of the history of CALL and its many acronyms, see Levy (1997).

[16]This same study is reported in Nagata & Swisher (1995). It focuses more on the idea of using intelligent feedback that provides metalinguistic information in a CALL environment as a means to consciousness-raising among language learners.

[17]A qualitative evaluation of the system was also conducted but will not be discussed in detail here, other than to say that the learners in both groups responded positively overall and that the error messages for the I-CALI feedback were ranked as being significantly preferred to those for the T-CALI feedback.

(2.11)   At a party your friend has asked if there are still Japanese drinks (left). Respond that no, the students went and drank all the Japanese beer (i.e., you were affected by the students drinking all the Japanese beer).

Friend: *Nihon no nomimono ga mada aru?*

Subjects in the traditional computer-aided language instruction (T-CALI) feedback group would be informed of missing, unexpected and incorrect words, whereas the intelligent computer-aided language instruction (I-CALI) feedback group would be informed of errors and told how to correct them using metalinguistic information about style, semantics, syntax and morphology. An example of a learner response to the question cited above, the correct answer and the feedback that a subject in the traditional group would have received are provided below (Nagata 1993, p.335):

(2.12)   **Correct answer:** *Uun, gakusee ni nihon no biiru o zenbu nomareta yo.*
          Incorrect answer: *Uun, gakusee ga nihon no biiru o zenbu nomaremasu.*[18]

    T-CALI feedback
    Read the following messages:
    <Particle error>
    - GA is not expected to be used here.
    - NI is missing.
    <Verbal predicate error>
    - NOMAREMASU is wrong.

In comparison, the messages that the I-CALI feedback group would have received for the same response would have read:

(2.13)   **I-CALI feedback**
    Read the following messages:
    <Particle error>
    - In your sentence, GAKUSEE is the 'subject' of the passive
    (the one that is affected by the action), but it should be the

---

[18]Nagata translates this as: *No, the students will be affected by someone's drinking all the Japanese beer.* She does not translate the correct sentence into English, however, a Japanese native-speaker colleague of mine proposed the following translation: *No, I am (negatively) affected by the students' drinking all the Japanese beer*.

'agent' of the passive (the one who performs the action and affects the subject). Use the particle NI to mark it.

<Verbal predicate error>

- The predicate you typed is in the imperfective form. Change it to perfective.

- Since you are talking with your friend and your friend is using the direct-style (casual style), use the direct-style for your response.

The post test results showed that the I-CALI feedback group significantly outperformed the T-CALI group. A tabulation of post test errors further revealed that the T-CALI group made 23.5% more errors on particles as compared to almost equal numbers of errors on vocabulary, verbs (i.e. tense, style, conjugations) and nominal modifiers. This pattern was maintained in the delayed post test. However, as the author concedes, speaking practice using the passive voice in the learners' drill class, which took place after the experiment, makes it difficult to attribute the delayed post test results solely to the I-CALI feedback. The study also found that, comparatively-speaking, more feedback messages resulted in fewer errors for the I-CALI feedback group on particles. Yet, more feedback for the intelligent group resulted in an equal number of errors on verbs with the traditional group[19] The author concludes that I-CALI feedback may be as good as T-CALI for improvement on word-level errors (i.e. verbal forms such as tense, style, negation and affirmation, conjugation and gerund forms) but better than T-CALI feedback for sentence-level errors (i.e. use of the particles *wa*, *ga*, *o*, *ni*, *e*, *de*, *no*, *kara* and *made*).

The author's conclusion is problematic based on three methodological oversights in her experiment. First of all, without having measured the learners' competence in producing the passive voice before the experimental treatment, a uniform level across groups could not be established. In fact, lower competence in verb forms in the I-CALI feedback group may explain why they needed more

---

[19]No mention is made of any correlation between amount of feedback and vocabulary or nominal modifier errors.

feedback than the T-CALI group. Secondly, as the immediate post test consti-
tuted the same task as the experimental materials, the results cannot be general-
ized beyond the task[20]. Thirdly, the strength of the author's claims with respect
to the effectiveness of the feedback are weakened by the fact that the learners
were able to refer back to the grammar notes on passivization that they read at
the beginning of each session at any point during the experimental treatment.
In light of these three points, the only conclusion that can be drawn from this
study is that in the context of the experimental task, I-CALI feedback tends to-
ward having an effect on improving the learners' command of particles in pas-
sive constructions relative to the effect of T-CALI.

Following the findings of her (1993) study, Nagata (1995) conducted a follow-
up experiment which was designed to pinpoint what aspect of I-CALI feedback
was effective in reducing errors in sentence production. Specifically, she wanted
to know whether the key element of I-CALI feedback was either the metalin-
guistic explanations of the grammatical and semantic functions of particles or
information about where they should be located in the sentence. To determine
this, she tested the same I-CALI feedback used in the previous study against an
enhanced form of T-CALI which not only told the learner which particles were
missing, unexpected or wrong, but also which verb phrases or noun phrases they
should be attached to. To facilitate comparison of the feedback between studies,
an enhanced T-CALI version of the feedback in example 2.12 is substituted be-
low. As mentioned, the I-CALI feedback would have remained the same (see
example 2.13):

(2.14)   **T-CALI feedback**
          Read the following messages:
          <Particle error>
          - GA is not expected to be used here.
          - NI is missing.
          - NI should be attached to GAKUSEE.
          <Verbal predicate error>
          - The form of the predicate you typed is wrong. Change the
          form.

---

[20]It could be that the delayed post test task was different from the experimental one. However,
no details of the test are given.

In light of the patterns that emerged from the previous study, the I-CALI feedback specifically addressed particle and verbal predicate errors. Both groups received the same messages for vocabulary errors.

Much of the methodology was replicated from the previous study. However, there were some important differences which should be mentioned. Firstly, the target structure was sentences in the active, rather than passive, voice using the particles *ga*, *wa*, *o*, *ni* and *de*[21]. Secondly, a pre test of ten gap-fill questions on these particles and two sentence production questions was conducted to measure within group competence in producing the target structure before the experimental treatment as well as to confirm a uniform level of proficiency[22]. An immediate post test was also carried out which included the pre test questions as well as an additional ten gap-fill and nine sentence production questions. In both the pre and post tests, the questions were similar to those asked in the experimental exercises. A delayed post test was not included in the design of the experiment because interfering variables, such as instruction, following the experimental treatment could not have been controlled[23]. Thirdly, there were fewer participants (18 as opposed to the 34 in the previous study) who were at a lower level of proficiency (taking a first-year, rather than a second year, Japanese course at an American university).

The overall results showed that both groups improved significantly from pre to post test[24] and that the I-CALI group outperformed the T-CALI group on the post test. A detailed inspection of the data revealed that the difference in post test scores between groups was due largely to particle errors, showing that metalinguistic explanations of the grammatical and semantic functions were what

---

[21]The functions of each particle are briefly explained in footnotes by Nagata (1993, p. 66) roughly as follows: *Ga* is a subject marker and it is also used to mark both the subject and object of stative predicates. *Wa* marks the topic and/or contrast of the sentence. *O* marks objects as well as the location through which an action moves. *Ni* marks the location where a non-activity referent is located, the direction toward which an action moves or the goal of an action. *De* marks the location where an activity takes place and the instrument by means of which an action occurs.

[22]The results of a midterm exam, rather than the pre test, were used to pair the learners and assign them to feedback groups on the basis of overall proficiency. No statistics are reported on the midterm scores. However, the results of t-tests on the pre test scores confirmed that there was no significant difference between the groups in their ability to produce the target structure.

[23]The author argues that this was less of a concern for the 1993 study, as the passive voice was not being covered in the course syllabus and the speaking practice that the learners received between the immediate and delayed post tests only took place during a single class.

[24]The author acknowledges, however, that claims regarding the contribution of the feedback to this improvement cannot be made without a control group.

contributed to improvement in particle production. This was not the case, however, for the verbal errors, as both feedback groups showed a similar number of errors on verb forms in the post test. To explain this result, the author proposes that the learners already knew the conjugations and therefore T-CALI feedback was sufficient for improvement to occur. She also suggests that the target conjugations were relatively simple and therefore detailed grammatical feedback was not necessary for improvement to be made. Whether or not the I-CALI feedback on verbal predicate errors was detailed is arguable, but it was not explanatory. Referring back to example 2.13[25], the learner is told simply that the imperfective form is wrong and that it should be changed to perfective. Why this should be the case is not explained. Furthermore, the feedback addresses an aspectual error rather than a conjugation one, as the author proposes. It seems more likely that lack of explanation is to blame for the difference in results for verbal predicate errors as compared to particle errors.

Nagata (1995) concludes that I-CALI feedback which explains grammatical and semantic function of Japanese particles is significantly more effective in developing sentence production skills than enhanced T-CALI feedback. She also hypothesizes that T-CALI systems may be suitable for vocabulary and conjugation exercises where word-level analysis is required while I-CALI may be superior for exercises of sentence production which require relatively complex syntactic analysis. With the inclusion of a pre test on the target structure, the results of this study are more robust than the previous one. However, it is still not possible to generalize the results beyond a single task.

Nagata (1996) is the third of her studies on I-CALI feedback. It is motivated by mixed findings in the CALL literature regarding the effectiveness of CALI as compared to more traditional forms of language instruction (e.g. language lab work, traditional classroom learning and pen and paper practice exercises). Focusing on the effects of feedback, Nagata (1996) sought to compare I-CALI and workbook instruction.

Nagata (1996) used the same level of proficiency measure (i.e. the mid-term exam results), target structures (i.e. sentence production using *ga*, *o*, *wa*, *ni* and *de*), I-CALI feedback and procedure as in the (1995) study. The pre and post tests were

---

[25]Note that verbal errors in this study refer to tense (imperfective or perfective) and negation but not style.

identical as well apart from one additional sentence production question on the pre test and one less on the post test for the (1996) study. The 26 subjects who participated in the study were all taking first-year Japanese classes. The workbook instruction group received exactly the same grammar notes and exercises as the I-CALI group, but on paper, and their exercises were corrected by the learners themselves using an answer sheet.

Effects of the experimental treatment on comprehension and oral production were also measured. Like the post test, the comprehension test was conducted immediately following the experimental treatment and repeated 23 days later to measure retention. It consisted of translating conversations in Japanese (13 sentences in total), of the type used in the experiment, into English. Oral production was first practiced throughout the three weeks following the experimental treatment for a total of one and a half hours. In pairs, the learners asked each other questions about what people did and where they went according to a series of pictures. Twenty-six days after the experiment, an oral test involving the same task was conducted with each subject individually. They were given the chance to ask and answer questions which required use of the target particles.

Analysis of the results of the mid-term exam and pre test confirmed that there were no significant differences between groups in level either overall or on written sentence production using particles. Comparison of the post test scores showed that the I-CALI group significantly outperformed the workbook group, demonstrating that I-CALI feedback is more effective than self-correction using answer sheets for developing grammatical competence in producing Japanese particles and sentences. Analysis was also conducted incorporating the data from Nagata's (1996) study in order to see if there was any difference between T-CALI and workbook instruction as well as I-CALI and workbook instruction. No significant difference was found between the T-CALI group and the workbook group; however, the I-CALI group in the (1996) study outperformed the workbook instruction group, consistent with the findings of the current study.

The results of the delayed post test between the I-CALI group and the workbook group[26] were significantly different. It is important to note that the oral practice described above took place between the immediate and delayed post tests.

---

[26]Recall that a delayed post test was not conducted in the (1995) study.

Furthermore, two days after the experimental period ended, the immediate post test was returned to the students and the answers to the test were explained to them. The immediate post test was subsequently reused as the delayed post test. The author interprets the result of the delayed post test to mean that "given the same follow-up instruction (i.e., feedback for the post-test and speaking practice), Nihongo-CALI is more effective than the workbook instruction in the long term" (Nagata 1996, p. 10). There are two problems with this conclusion. Firstly, since the pre and post test questions and the experimental treatment questions followed a similar format, the results cannot be generalized beyond the task. Secondly, even though the follow-up instruction was the same for both groups, it is not certain that their *effect* was the same on both groups. It could have been the case, for instance, that there was an effect of combining I-CALI feedback specifically and speaking practice that would not have occurred as a result of I-CALI feedback alone. Therefore, without controlling for the additional factor of follow-up instruction, any long term effects of the experiment cannot be attributed solely to the I-CALI feedback. The same criticism holds for the results of the oral test which showed a significant difference in favour of the I-CALI group following the same speaking practice for both groups. Furthermore, since the pre test only measured written production, there is no confirmation that the two groups were at a uniform level in their oral production. Therefore, the possibility that the results are due to the I-CALI group having higher oral competence before exposure to feedback cannot be excluded.

The results of the two comprehension tests (immediate and delayed) showed no difference between the I-CALI and workbook groups. The author attributes this to the fact that understanding the grammatical and semantic functions of the particles was not a necessary pre-condition for the learners to be able to assign an accurate interpretation to the sentences. From the sample question provided in the article, this seems to have indeed been the case. Given the sentence *Kinokuniya de omosiroi hon o kaimasita* (I bought an interesting book in Kinokuniya), it is clear that if the learner knows that *Kinokuniya* is the name of a book store, *omosiroi hon* means "interesting book" and *kaimasita* means "bought", there is a single interpretation of the sentence which can be arrived at by using logic alone. The author then goes on to conclude that this result, in association with the production data, confirms that sentence production requires more syntactic

processing than comprehension. She also suggests that a comprehension test in isolation may be insufficient for accurately measuring a learner's grammatical skill. While these two points may be valid, the results of the comprehension test are not evidence of them. Rather, they show that the comprehension questions did not tap into the knowledge that the author wanted to test and, consequently, are not relevant to the study.

Nagata (1997a) describes the development of *BANZAI* a replacement system for *Nihongo-CALI*. In the same article, the author reports on a replication of her (1996) experiment which was conducted using *BANZAI* and tested the effect of inductive versus deductive feedback on the Japanese particles *ga, o, wa, ni* and *de*.

The deductive feedback identifies the learners' error, informs them of the violated rule and supplies the correct answer. The inductive feedback does likewise, except that instead of referring to a rule, two (or, in some cases, three) examples using the target particle are given. The example below, modified from Nagata (1997a, p. 524), illustrates the deductive feedback, inductive feedback, student answer and correct answer to this question: *Your tutor has asked what you did last night. Respond that you wrote a letter to a friend in Japanese.*:

(2.15)   **Correct answer:** *Tomodati ni nihongo de tegami o kakimasita.*[27]
Student answer: *Tomodati ga nihongo de tegami o kakimasita.*
Deductive feedback:
You used the particle GA to mark TOMODATI as though it had the role SUBJECT (the one who performs the action). But the correct role is GOAL (the goal of the action). Use NI to mark it.
Inductive feedback:
You used the particle GA to mark TOMODATI, but the correct particle is NI. The following examples show how the particles NI and GA are used.
Particle NI:
  1. Tanaka-san ni kakimasu. '(I) will write to Mr./Ms. Tanaka.'
  2. Tomodati ni misemasu. '(I) will show (it) to my friend.'
Particle GA:
  1. Tanaka-san ga ikimasu. 'Mr/Ms Tanaka will go.'

---

[27] friend-*Dat.* Japanese-in letter-*Acc.* (I) wrote (I wrote a letter to a friend in Japanese.)

2. Tomodati ga tukurimasita. 'My friend made (it).'

The results of the post test were divided into scores for the fill-in-the-blank questions and scores for the sentence production questions. Both were found to be significantly different in favour of the deductive feedback group, indicating that deductive feedback was more helpful than inductive feedback for learning Japanese particles and sentences. This was also the case for the results of the delayed post test, showing that deductive feedback was superior to inductive feedback in helping learners to produce particles over three weeks. The results of the two comprehension tests were not significantly different, but the oral results were, again in favour of the deductive feedback group. The author interprets the findings of the oral test to mean that deductive feedback was more effective than inductive feedback for the oral tasks used in the study.

Most of the criticism lodged against the claims made in the (1996) study are applicable to this one. Specifically, the relative effectiveness of the deductive feedback cannot be generalized beyond the task[28], claims regarding retention being attributable solely to the feedback cannot be maintained due to the speaking practice which occurred between the immediate and delayed post tests and the results of the comprehension test should have been discounted on grounds that the test did not assess the knowledge under scrutiny.

The four studies described thus far used either *Nihongo-CALI* or *BANZAI*, ILTSs with NLP capability. Nagata (1997b) opts for a system that does not carry out any linguistic analysis but can still provide intelligent feedback about the Japanese particles *ni*, *o*, *ga*, *de*, *wa* and *to*. In this study, a preliminary experiment was conducted to ascertain what strategies learners use when assigning particles. They were provided with a context and had to fill the gaps in a sentence that was part of a written conversational exchange with the correct particles. They were then asked to to justify their choice of particle. The study found that when asked to explain why they chose a particular particle, the learners either stated a metalinguistic rule or provided an equivalent in English. Therefore, the main study compared the relative effectiveness of these two strategies as feedback.

---

[28]Nagata (1997a, p. 528) seems to acknowledge this with respect to the oral results, as she states "...deductive feedback was more effective than inductive feedback **for the oral conversation tasks involved in the study**." (emphasis added). This is not the case in her interpretation of either the comprehension or the written results.

Fourteen learners enrolled in a second-year university Japanese course partici-
pated in the experiment. The preliminary study served as a means for establish-
ing level on particle knowledge and was used to divide the learners into groups
of equal proficiency. The experimental task was the same as the one used for
the preliminary study. Upon making an error, learners in both groups were in-
vited to try again. On a failed second attempt, feedback would be provided to
the learner. For both groups, this meant being explicitly told that the particle
was wrong and providing the correct particle. This was followed by an expla-
nation of the grammatical and semantic functions of the correct particle, in the
case of the metalinguistic feedback group, and a translation of the noun, verb
or prepositional phrase associated with the erroneous particle, for the transla-
tion group. The metalinguistic and translation feedback were also provided in
response to correct particles in order to confirm that the reason for the learners'
particle choice was accurate. The learners participated in three computer ses-
sions of approximately one hour each and completed a total of 64 exercises in the
space of a week.

The results of an immediate post test, using questions similar to those in the ex-
perimental treatment, were compared between the two groups[29] and showed a
statistically significant difference in favour of the metalinguistic feedback group.
The author concludes that metalinguistic feedback is more effective than transla-
tion feedback in helping learners to master complex grammatical structures. As
the author concedes, the robustness of these findings are limited by the absence
of a control group, pre test and delayed post test.

The feedback used in the ILTS designed by Ferreira-Cabrera (2003) was informed
by two observational studies, one set in a classroom and the other in a tuto-
rial, which were conducted as part of a research project investigating feedback
strategies for second language teaching and their implications for the design of
ILTSs. Following Chi et al. (2001), Ferreira-Cabrera (2003) categorized feedback
types as either Prompting Answer Strategies (PAS) or Giving Answer Strategies
(GAS). For the purposes of her study, PAS were defined as metalinguistic cues,
clarification requests and elicitation and GAS were identified as repetition of the
error with rising intonation, recast, provision of the correct answer and explicit

---

[29]As the preliminary study used fewer particles than the experimental treatment and the im-
mediate post test, it could not be used in the analysis as a pre test.

correction. Having found in the observational studies that PAS resulted in more frequent repair than GAS for grammar and vocabulary errors[30], Ferreira-Cabrera (2003) sought to determine experimentally whether or not PAS would be more effective than GAS for rectifying grammar errors as well as whether both would be more effective as compared to no feedback.

The experiment used the Spanish subjunctive as its target structure. Specifically, the present and imperfect subjunctive following expressions of hope/desire, advice/suggestion, personal preference and possibility were tested. Twenty-four subjects participated, 22 of whom spoke English as a first language. The remaining two were native speakers of French and Portuguese, respectively. As native speakers of a language different from the rest of the subjects and, particularly, of languages that belong to the same family as Spanish, these two subjects should have been excluded from the experiment in order to control sufficiently for the independent variable.

The pre test, post test and experimental tasks were all centred around correspondence and discussion about a trip to Chile. For instance, the pre test consisted of a letter sent by Jill to her Chilean friend María explaining what parts of the country she and her friends were interested in seeing on their holiday. The interaction continued throughout the three experimental tasks and finished with a post test which was an email from Jill thanking María for her help and signaling the end of the correspondence. Within this context, the learners were asked to carry out both production and comprehension tasks. The production task consisted of completing blanks in the correspondence with the correct verb form and the comprehension task required that they choose the correct verb from a choice of four forms.

Of the PAS identified in the observational studies, metalinguistic prompts were selected for experimental testing. These were essentially statements of the rule that the learners violated in their responses. For instance, consider the following incorrect response (Ferreira-Cabrera 2003, p. 163). Here the learner has provided the imperfect indicative form, *llovía*, instead of the target-like present subjunctive, *llueva*:

---

[30]Although excluded from the ILTS experiment, pronunciation errors were taken into account in the observational studies. The author found that GAS were the only strategies used by teachers and that they resulted in repair.

(2.16)   Incorrect response:

Dudo mucho que *llovía* en el norte de Chile, porque es desierto.

*I doubt it **rained** in northern Chile because it is desert.*

In reply, the metalinguistic prompt would read:

(2.17)   PAS (first error):

"Con las expresiones o verbos que expresan la idea de duda es adecuado usar el modo subjuntivo."

(With verbs expressing the idea of doubt, the subjunctive mood should be used.)

In the event that the learner makes a second error in mood[31], a similar metalinguistic prompt would be provided:

(2.18)   PAS (second error):

"Fíjate que con los verbos que expresan duda es adecuado usar el modo subjuntivo."

(Note that with verbs expressing doubt, it is appropriate to use the subjunctive mood.)

Two types of GAS were chosen for experimental testing. The first was repetition of the error with a question mark to draw attention and was provided in response to the learners' first error. In reply to the learner response in example 1.16 cited above, the feedback would be as follows:

(2.19)   GAS (first error):

llovía?

The second was explicit correction which entailed presenting the correct form and signaling it as such. For the same example, the feedback would have been thus:

(2.20)   GAS (second error):

llueva (correction)

---

[31]Note that the ILTS prioritizes mood errors over tense errors. To address a response to example 1.16 where the mood is appropriate but the tense is not (i.e. the imperfect subjunctive has been used instead of the present subjunctive), the learner would receive a feedback message such as: "Tienes que considerar la secuencia de los tiempos. 'Dudo' corresponde a presente..." (*You have to consider the tense. 'Dudo' is a present tense.*)

T-tests on the pre and post test scores showed significant improvement for all groups. However, comparison of gain scores (i.e. the difference between the post and pre test scores) revealed a significant effect in favour of the PAS group, showing that metalinguistic prompts were more effective than both repetition followed by explicit correction and no feedback. Another interesting finding of the experiment was that throughout the experimental tasks both the PAS and GAS groups maintained similar scores which showed improvement, but the GAS group's average post test score was lower than their scores on two of the three experimental exercises. The author suggests that while GAS had an effect when they were being provided to the learner, this was not the case when the feedback was removed.

Unfortunately, the claims made about the effectiveness of PAS are limited by how improvement was measured. From close inspection of the pre test, post test and experimental questions, it is clear that many of the verbs introducing a subordinate clause in the subjunctive were repeated several times in the course of the experiment. For instance, *gustar* was used to express personal preference in the pre test[32] and sets one[33], two[34] and three[35] of the experimental exercises, meaning that the learners had answered three questions using the same verb before being asked to do so again on the post test. This was the case for the expressions of hope/desire and personal preference which accounted for 50% of the target-structure questions on the post test. As for the remaining questions, the same verbs had been seen less than three times, and in some cases only once, by the learners during the experiment. It remains unclear, then, whether PAS enabled the learners only to memorize particular lexical items or whether they helped the learners to form generalizations concerning target-like usage of the subjunctive.

---

[32]Pre test, Question 9: ***Nos gustaría que*** *tú nos [acompañabas / acompañaras / acompañes / acompañas] a conocer la región del norte de Chile.*

[33]Set 1, Question 10: ***Me gustaría mucho que*** *nosotros [pasamos / pasabamos / pasemos / pasaramos] algunos días recorriendo Santiago.*

[34]Set 2, Question 10: ***Me gustaría que*** *[conozcas / conocías / conocieras / conoces] las hermosas figuras de oro del Museo Arqueológico del Padre Le Paige.*

[35]Set 3, Question 6: ***Me gustaría que*** *tú nos [acompañes / acompañabas / acompañaras / acompañas] en nuestro viaje al Norte, pero entiendo que estás muy ocupada en tu trabajo.*

2.4.1 Evaluative Studies of ICALL Feedback Summary

Based on the literature amassed to date, the potential of ICALL feedback remains promising. It cannot be denied that the absence of control groups in all of the studies reviewed here by Nagata weakens her claims with respect to the effectiveness of feedback. In addition, flaws in experimental design pose problems for suggesting that there are any long term effects of ICALL feedback. However, the statistical analyses in all of her experiments consistently showed that, immediately following the experiments, ICALL feedback was significantly more effective than the other types of feedback that were tested. PAS in Ferreira-Cabrera's work were also found to be significantly more effective than GAS and no feedback, even though the precise interpretation of the results is problematic. In order to answer the question of whether ICALL feedback is effective, much more research that explores other second languages and different target structures based on sound methodological design is necessary.

Effective ICALL feedback, as it was operationalized in the Nagata and Ferreira-Cabrera studies, responded to errors by giving a metalinguistic explanation in the form of a rule, a type of feedback referred to as deductive in Section 2.3.4[36]. The motivation for testing deductive feedback over other types of feedback in these studies originated from several sources. With a view to identifying feedback strategies appropriate for an ILTS, Ferreira-Cabrera (2003) used what she observed in language teaching classrooms to motivate her choice of feedback. Nagata (1997a) drew from her own study about what strategies learners themselves used in producing particles in Japanese. The over-arching motivation for the earlier work by Nagata (1993, 1995, 1996) was to demonstrate the value of ICALL over traditional CALL and classroom instruction. Like the work that is driven by NLP, reviewed in Section 2.2.1, there was no particular reason for using deductive feedback other than it made use of all the information about the errors that had been gathered during the parse. All of this work has employed pragmatic justification for choosing deductive feedback. Nagata (1997b), on the other hand, motivated her choice of feedback from the theoretical debate in the

---

[36]In the psycholinguistic studies reviewed in that section, the deductive feedback was explicit, meaning that the learners were told that their response was wrong. This was also the case in the Nagata (1993, 1995, 1996, 1997a,b) studies. However, Ferreira-Cabrera (2003) chose implicit deductive feedback whereby the rule was given without informing the learners directly that they were wrong.

applied linguistics and second language acquisition literatures about deductive versus inductive language teaching. Work in the second language acquisition literature on feedback is another area which has yet to be tapped as a source of motivation for testing particular feedback types.

Empirical work on feedback in both the second language acquisition and ICALL literatures has shown an effect for deductive feedback, whether implicit or explicit.[37] However, as discussed in Section 2.3.4, not every area of grammar can be readily embodied as rules such as those used in deductive feedback. Thus, we propose explicit inductive feedback as a new candidate for testing with the idea that it might maintain the explanatory benefit of deductive feedback without compromising the accurate representation of the target structure. Furthermore, it would allow for much needed expansion of the small base of evaluative studies of ICALL feedback. Moreover, it would fill a gap in the second language acquisition literature whose current focus on feedback is oriented towards testing the predictions of the Interaction Hypothesis.

2.5 Processing Instruction

Another research question in the instructed second language acquisition literature, as identified by Norris & Ortega (2000), compares the effect of traditional grammar explanation and practice to an approach to grammar teaching which encourages learners to process input in psycholinguistically relevant ways. The research that this question refers to is that of Processing Instruction (PI) and the input processing model on which it is based. As the name indicates, PI is a type of instruction, not feedback. However, the essence of the approach has the potential to also apply to feedback and, on the basis of the success of PI, would be worth investigating. With this in mind, we will first turn our attention to briefly reviewing the details of Processing Instruction, its theoretical underpinning and the results of studies that have tested its effectiveness.

---

[37]Support has been found for both implicit and explicit deductive feedback but the findings for the explicit variety are more robust simply because it has been tested in more studies. For this reason, we chose to make our inductive feedback explicit rather than implicit.

PI is an approach to grammar teaching developed by Bill VanPatten (1993, 1996). Seeking to bridge the gap between second language teaching and second language acquisition theory and research, VanPatten (1993, 1996, 2002a,b, 2004a) motivates PI by deriving it from a model, proposed by him, of how second language learners process input. The model consists of a series of principles and their subparts which express a number of constraints on how learners make connections between form and meaning while processing input.[38] To elucidate the findings of studies that have tested PI, the principles of input processing, in their most recent formulation, are stated below (VanPatten 2004a, p. 14 and p. 18):

Principle 1. The Primacy of Meaning Principle. Learners process input for meaning before they process it for form.

> **Principle 1a. The Primacy of Content Words Principle.** Learners process content words in the input before anything else.
>
> **Principle 1b. The Lexical Preference Principle.** Learners will tend to rely on lexical items as opposed to grammatical form to get meaning when both encode the same semantic information.
>
> **Principle 1c. The Preference for Nonredundancy Principle.** Learners are more likely to process nonredundant meaningful grammatical form before they process redundant meaningful forms.
>
> **Principle 1d. The Meaning-Before-Nonmeaning Principle.** Learners are more likely to process meaningful grammatical forms before nonmeaningful forms irrespective of redundancy.
>
> **Principle 1e. The Availability of Resources Principle.** For learners to process either redundant meaningful grammatical forms or nonmeaningful forms, the processing of overall sentential meaning must not drain available processing resources.
>
> **Principle 1f. The Sentence Location Principle.** Learners tend to process items in sentence initial position before those in final position and those in medial position.

Principle 2. The First Noun Principle. Learners tend to process the first noun or pronoun they encounter in a sentence as the subject/agent.

---

[38]Note that VanPatten (2004b) considers input processing to consist of two sub-processes: establishing form-meaning connections and parsing.

Principle 2a. The Lexical Semantics Principle. Learners may rely on lexical semantics, where possible, instead of word order to interpret sentences.

Principle 2b. The Event Probabilities Principle. Learners may rely on event probabilities, where possible, instead of word order to interpret sentences.

Principle 2c. The Contextual Constraint Principle. Learners may rely less on the First Noun Principle if preceding context constrains the possible interpretation of a clause or sentence.

The model of input processing that these principles embody has met with criticism, particularly from DeKeyser et al. (2002). They have two major objections to the model. The first is that the model of attention adopted by VanPatten, which is that learners have limited attentional capacity, is outdated. The second is that the model is not consistent with the findings of current research on sentence processing. For instance, principle 1f (the Sentence Location Principle) claims that learners process items in initial, followed by final, followed by medial position while current models of sentence processing clearly show that parsing relies heavily on, or is at least constrained by, structural information (DeKeyser et al. 2002) and proceeds incrementally (Carroll 2004). Whether or not principle 2 (the First Noun Principle) applies to most language learners is also questionable. Carroll (2004) suggests that processing the first noun as a subject is unlikely to hold for speakers of topic prominent languages. As for agentivity, she argues that it is clear from work that has compared cross-linguistic processing strategies that not all speakers assign the first noun the semantic role of agent.

Despite the problems with the underlying model of input processing, PI has exhibited substantial success as a type of focus on form instruction. We turn now to a brief description of how PI works, followed by a summary of related empirical work.

2.5.2 *The Mechanics of Processing Instruction*

PI involves three steps: firstly, the learners are given an explanation of the target structure; secondly, they are informed of processing strategies that are predicted to cause problems in processing the target structure and thirdly, they practice

mapping the meaning of the target structure with its form by way of "structured input activities". For instance, suppose, as VanPatten & Cadierno (1993) did, that the target structure was object pronouns in Spanish. In their experiment, instead of presenting the complete paradigm of object pronouns, an explanation was given by person (e.g. the third person pronouns *lo*, *la*, *los* and *las* were explained in a chunk). A series of examples in Spanish with English translations contrasting subject and object pronouns were used to illustrate the form, meaning and syntax of third person pronouns. Principle 2 of VanPatten's input processing model predicts that learners "...tend to process the first noun or pronoun they encounter in a sentence as the subject/agent." (VanPatten 2004a, p. 18). Since Spanish is a pro-drop language and object pronouns typically occur pre-verbally[39], it is this processing strategy which may cause learners to interpret object pronouns as subject pronouns/agents. Accordingly, in step two of processing instruction on object pronouns, the learner is warned against applying Principle 2 when processing input. This has been expressed by VanPatten & Cadierno (1993), as cited in VanPatten (1996, p. 65), as follows: *Keep in mind that Spanish has flexible word order and doesn't necessarily follow subject-verb-object order like English.* It has also been articulated more specifically by Cadierno (1992) and VanPatten et al. (1992), as cited by VanPatten (1996, p. 73), in the following manner: *Keep in mind that Spanish does not follow a rigid subject-verb-object word order and that object pronouns may go before a conjugated verb or at the end of an infinitive.*

Structured input activities require that the learners process the second language, rather than produce it. The activities are structured in that in order to complete the tasks successfully, the learners must match the meaning of the target structure to its form, thereby rectifying the hypothesized faulty input processing strategy. A structured input activity is provided below (VanPatten 1996, p. 65):

(2.21)   Structured input activity for Spanish object pronouns

      1. *Mi hermana me llama frecuentemente.*[40]
      Who calls whom?
      a. I call my sister   b. My sister calls me.

---

[39]This is always the case, except in sentences involving an infinitive, in which case the pronoun can optionally appear as a suffix on the infinitive.

[40]My sister-*Nom.* me-*Acc.* calls frequently

2. *¿Te escriben tus padres?*[41]

Who writes to whom?

a. Do you write to your parents?   b. Do your parents write to you?

3. *No nos escuchan los padres*[42]

Who doesn't listen to whom?

a. Parents don't listen to us.   b. We don't listen to parents.

4. Me conocen bien mis hermanos.[43]

Who knows whom well?

a. My siblings know me.   b. I know my siblings.

In order to successfully answer the question of who did what to whom, the learner must be able to differentiate between subject and object. Furthermore, they cannot rely on word order to make this distinction, as the order varies between SOV and OVS, depending on the question (e.g. question 1 above is SOV whereas question 4 is OVS). Consequently, the structured input activity verifies that the learners are processing the form correctly and, in the event they are not, makes the correct answer obvious by process of elimination.

2.5.3 Empirical Work on Processing Instruction

PI has been tested empirically on a number of target structures. These include object pronouns and word order (Sanz 1997, 2004, VanPatten & Cadierno 1993, VanPatten & Fernández 2004, VanPatten & Oikkenon 1996, VanPatten & Sanz 1995), the conjugations of the preterite tense (Cadierno 1995), the copulas *ser* and *estar* (Cheng 2002, 2004) and the subjunctive (Farley 2001a,b, 2004a,b) in Spanish; the *faire causitif* construction (VanPatten & Wong 2004) and indefinite articles in affirmative versus negative declarative sentences (e.g. *J'ai <u>une</u> pomme*[44] vs *Je n'ai pas <u>de</u> pomme*[45] in French (Wong 2004) and the conjugations of the future tense in Italian (Benati 2001, 2004).

---

[41]You-*Dat.* write your parents-*Nom.*

[42]*Neg.* us-*Dat.* listen the parents-*Nom.*

[43]Me-*Acc.* know well my siblings-*Nom.*

[44]I have an apple.

[45]I *Neg.* have *Neg.* an apple.

In studies comparing the effects of PI and Traditional Instruction (TI)[46], PI has been consistently superior to TI in helping learners in comprehension and as good as TI in improving production of the target structures[47] (Cadierno 1995, VanPatten & Cadierno 1993, VanPatten & Wong 2004). Further studies have shown that the effects of PI are generalizable beyond sentence-level tasks (i.e. to narration tasks) (Sanz 2004, VanPatten & Sanz 1995) and retained in the long-term (i.e. over eight months)[48] (VanPatten & Fernández 2004). The remaining work (Benati 2004, VanPatten & Oikkenon 1996, Wong 2004) has explored PI *per se* with a view to identifying what aspect – the explicit information (EI)[49], the structured input activities (SI) or both[50] (i.e. PI) – of the approach accounts for its effectiveness. In all cases, learners who received SI or PI about the target structure made significant improvements on the interpretation task as compared to the EI group and the controls. This was also the case for the learners in the Benati (2004) study on the production task. Similarly, in the VanPatten & Oikkenon (1996) study, the SI and PI groups showed higher gains on the production task between pre and post test, although they did not reach significance. Wong (2004) found that the PI and SI significantly outperformed the controls, but only the PI group outperformed the EI group on the production task. The general consensus is that the structured input activities are the crucial component of PI for comprehension. There seems to be a similar pattern with respect to production; however, the results are less robust.

Not all studies have shown superior results for PI. Collentine (1998) and Farley (2004a) found that both PI and meaning-based output instruction (MOI) groups

---

[46]Traditional Instruction is defined as that in which learners receive rules about how a structure works, followed by output exercises which range from mechanical to meaningful to communicative. For the study conducted by VanPatten & Cadierno (1993) on object pronouns and word order in Spanish, for example, the traditional instruction consisted of presenting the learners with the full paradigm of object pronouns, explaining what they are and where they are placed in a sentence and leading the learners through a number of oral and written mechanical, meaningful and communicative exercises.

[47]This study was conducted as a replication of Allen (2000) which, in turn, was a replication of VanPatten & Cadierno (1993). Allen found that both the PI and TI groups improved on the interpretation task but that the TI group outperformed the PI group on the production task. The discrepancy in the findings was due to differences in the pre/post test and experimental tasks.

[48]It should be noted, however, that the claims made by VanPatten & Fernández (2004) are weakened by the fact that their study did not include a control group.

[49]Note this acronym is unrelated to the term explicit inductive (EI) which we use to describe one of our feedback groups.

[50]Note that the statement of the faulty processing strategy was included in both the explicit information and in the structured input activities.

improved on the uses of the Spanish subjunctive in relative clauses and to express doubt, respectively, corroborating the claim made by DeKeyser & Sokalski (1996) that there is a qualitative difference between traditional instruction and PI, namely that TI promotes attention to form, whereas PI encourages attention to form *and* meaning. Cheng (2002, 2004) acknowledges that this also may explain why the results of her studies on the lexical-aspectual distinction between *ser* and *estar* did not categorically favour PI. There has also been some suggestion that the complexity of the target structure may be an intervening factor (Collentine 2002).

Although PI has not been shown to be superior to other types of instruction in all cases, it has in many. Furthermore, it has consistently promoted improvement in production comparable to that achieved by output-based instruction, despite the fact that it is an input-based approach which does not provide any opportunity for output practice. What appears to be the crucial factor in its success is the structured input activities, in which the learners are given a sentence or scenario and are asked to match it with one of two proposed interpretations. This could be adapted as a type of feedback, whereby in response to an error, the learners are presented with a correct interpretation and an interpretation for the learner's answer and are required to match their own response with its interpretation. We will refer to this type of feedback as Input Processing (IP) feedback. How exactly it would work for aspect will be described in Section 5.4.2.

2.6   Chapter Summary

The foregoing review has revealed a need for further research on feedback in both the ICALL and second language acquisition literatures. Previous work in ICALL has focused on developing and describing prototype systems purporting innovations in NLP, pedagogy or feedback. Of the research that has concentrated on feedback, only work by Ferreira-Cabrera (2003) and Nagata (1993, 1995, 1996, 1997a,b, 1998), Nagata & Swisher (1995) has attempted to quantitatively establish the usefulness of ILTSs or the effectiveness of the feedback they deliver. Furthermore, little research (Heift 2001, Nagata 1997b) has tested theoretically-motivated feedback and none, to our knowledge, has exploited the literature on feedback in second language acquisition. Seeing that work on feedback in the

second language acquisition literature is currently dominated by testing the predictions of the Interaction Hypothesis, we also identified a need for expansion of promising work outside of this framework.

The study that we are proposing, which will be described in detail in chapter 5, will address issues in both fields by testing the effectiveness of two, as yet untested, types of feedback. Both feedback types were borne out of previous work in ICALL and second language acquisition. The first was EI feedback which was motivated by the significant positive results for explicit deductive feedback in both literatures. We reasoned that EI feedback would maintain the explanatory benefit of explicit deductive feedback but that providing specific meanings, rather than rules, would more accurately represent complex areas of grammar, such as aspect. The second type of feedback was IP. Following the success of Processing Instruction (Cadierno 1995, Sanz 2004, VanPatten & Cadierno 1993, VanPatten & Fernández 2004, VanPatten & Sanz 1995, VanPatten & Wong 2004) and of structured input activities in particular (Benati 2004, VanPatten & Oikkenon 1996, Wong 2004), we decided to incorporate the idea behind structured input activities into a type of feedback. As such, IP feedback will encourage target-like mapping of aspectual form to aspectual meaning by requiring that learners assign a specific meaning to the form of their response.

CHAPTER 3

Aspect

3.1   Introduction

The first section of this chapter will be dedicated to presenting current theoretical terminology and concepts fundamental to the understanding and discussion of aspect. Section 3.3 will contrast the aspectual systems of French and English and Section 3.4 will explore the notion of aspectual coercion in both languages. In light of the discussion in Sections 3.3 and 3.4, Section 3.5 will identify the dimensions of aspect that might pose difficulties to learners on the basis of language transfer. They will constitute the main target structures which the feedback for the experiment will address (see also Section 5.6.1 for further details).

3.2   The Semantics of Aspect

In contrast to tense, which relates the time of a situation to some other time, usually the moment of speaking (Comrie 1976), aspect conveys a temporal perspective on a situation. This temporal perspective is a focus either on all or part of the situation (Smith 1991). While there are many competing analyses of aspect in the temporal semantics literature (see Binnick 1991 for an overview), semanticists seem to agree that the aspectual system of a language consists of two basic components: **lexical aspect**[1] and **grammatical aspect**.

---

[1]Depending on the theoretical framework, lexical aspect can be referred to as situation(al) aspect, inherent (lexical) aspect, situation type (Smith 1991), eventuality description (de Swart 1998) or Aktionsart. All of them refer to the same concept, with the exception of Aktionsart which is argued by some (see Comrie (1976) and Binnick (1991)) as not being equivalent to lexical

Lexical aspect refers to the temporal properties of a given situation which are inherent in a verb and its arguments. Attempts to divide lexical aspect into categories date back to Aristotle. In fact, it was he who inspired the four-way classification proposed by Vendler (1967) which is the most commonly used in the second language acquisition literature[2]. The four categories are *states*, *activities*, *accomplishments* and *achievements* and are defined by Smith (1991, p. 28) as follows[3,4]:

States are static, with no dynamics and no internal structure; they have duration of at least a moment. Examples: [know the answer], [be in Athens].

Activities are durative atelic events; they have homogenous successive stages and an arbitrary final point, e.g. [walk in the park], [laugh].

Accomplishments are durative telic events: they are complex, consisting of a process of successive stages and a natural final point. The outcome is a change of state, e.g. [build a house], [walk to school].

Achievements are instantaneous changes of state, with an outcome of a new state, e.g. [reach the top], [win a race].

These four categories are defined according to the temporal properties of dynamism, durativity and telicity. Dynamism contrasts events and states. States are

---

aspect. We will maintain the term *lexical aspect* for the sake of simplicity and because it is the most commonly used in the second language acquisition literature.

[2]We acknowledge that many semanticists argue for conflating accomplishments and achievements into a single category of telics, reducing Vendler's four classes to three (Verkuyl 1989). For instance, Bach (1986), Mourelatos (1978) and Piñón (1995) propose the categories of states, processes and events, Taylor (1977) and Dowty (1986) advocate states, energeia and kinesis and Verkuyl (1993) opts for statives, activities and telics. The justification for a three-way classification is that the feature [±punctual], which is supposed to distinguish achievements from accomplishments, actually does not: many achievements do have some duration (Dowty 1986). For instance, *die* is classified as an achievement but can be seen as having multiple stages, as in *John was dying when the doctor arrived* (Dowty 1986, p. 42).

[3]These four lexical aspectual classes have been articulated by many researchers. We chose to cite Carlota Shipman Smith's version because it seemed to be the most clearly expressed. We refer the reader to Andersen & Shirai (1995, p. 744) as well as to Vendler (1967, p. 106) himself for other articulations of the four categories.

[4]Smith (1991) also includes semelfactives in her classification. They have not been explained here because they have not been considered a separate category in the acquisition studies.

Situations	Static	Durative	Telic
States	[+]	[+]	-
Activity	[-]	[+]	[-]
Accomplishment	[-]	[+]	[+]
Achievement	[-]	[-]	[+]

Table 3.1: Four Classes of Lexical Aspect by Feature

"unchanging throughout their duration" (Binnick 1991, p. 183), whereas events are not. In other words, all the stages of a state, such as [be in Athens], are the same, while those of an event, such as [build a house], are not. Telicity refers to an event which has a natural endpoint that is signalled by the occurrence of a change of state. It distinguishes, for instance, activities from accomplishments. While an activity, such as [walk in the park], does not have a natural endpoint, an accomplishment, such as [build a house], does. Durativity contrasts events that are durative to those which are instantaneous. An accomplishment, such as [build a house], is differentiated from an achievement, such as [win a race], by the fact that winning a race is instantaneous while building a house is not. Table 3.1 shows a feature analysis of the four classes of lexical aspect (adapted from Smith (1991, p. 30) excluding semelfactives)[5]:

Diagnostic tests for distinguishing members of each category have been developed by semanticists[6] and adapted for use in second language research. While diagnostic tests exist for several languages in the second language acquisition literature (see Bardovi-Harlig & Bergström (1996) and Bergström (1995, 1997) for French, Hasbún (1995) for Spanish and Shirai (1995) for Japanese), we will present an example of one for English[7] (Andersen & Shirai 1995, p. 749). Typically, tests are not ordered; however the following ones happen to be:

(3.1)   **Step 1**: State or nonstate

Does it have a habitual interpretation in simple present tense?

If no → State (e.g. *I love you*)

If yes → Nonstate (e.g. *I eat bread*) → Go to Step 2

---

[5]We note that Smith does not put the telic feature for states in square brackets because she considers it no longer relevant for situations that are [+static]

[6]For a summary of the tests developed by Dowty (1979) and Freed (1979), see Binnick (1991, pp. 173-178).

[7]For other sources of diagnostic tests for English, see Bardovi-Harlig & Bergström (1996), Bardovi-Harlig (1998), Robison (1990, 1995) and Shirai (1991).

Step 2: Activity or nonactivity

Does 'X is Ving' entail 'X has Ved' without an iterative/habitual meaning? In other words, if you stop in the middle of Ving, have you done the act of V?

> If yes → Activity (e.g. *run*)
>
> If no → Nonactivity (e.g. *run a mile*) → Go to Step 3

Step 3: Accomplishment or achievement

[If test (a) does not work, apply test (b) and possibly (c).]

> (a) If 'X Ved in Y time (e.g. *10 minutes*)', then 'X was Ving during that time.'
>
> > If yes → Accomplishment (e.g. *He painted a picture*)
> >
> > If no → Achievement (e.g. *He noticed a picture*)
>
> (b) Is there ambiguity with *almost*?
>
> > If yes → Accomplishment (e.g. *He almost painted a picture* has two readings: he almost started to paint a picture/he almost finished painting a picture)
> >
> > If no → Achievement (e.g. *He almost noticed a picture* has only one reading)
>
> (c) 'X will VP in Y time (e.g. *10 minutes*)'= 'X will VP after Y time.'
>
> > If no → Accomplishment (e.g. *He will paint a picture in an hour* is different from *He will paint a picture after an hour*, because the former can mean that he will spend an hour painting a picture, but the latter does not.)[8]
> >
> > If yes → Achievement (e.g. *He will start singing in two minutes* can have only one reading, which is the same as in *he will start singing after two minutes*, with no other reading possible.)

---

[8]We believe that the use of the future tense here is problematic because it renders *He will paint a picture in an hour* ambiguous. While *He will paint a picture in an hour* can mean that he will spend an hour painting a picture, it can also mean that an hour will elapse and then he will paint a picture. Using the past tense dispels this ambiguity because *He painted a picture in an hour* can only mean the painting of the picture took an hour.

Even though the examples are marked for tense and aspect, Andersen & Shirai (1995) specify that the tests should be applied to verbs and their arguments without any such markings (i.e. [she sing a song] rather than *she sang a song* or *she was singing a song*).

As a final point, we wish to clarify that lexical aspect is determined by the verb and its arguments, including the subject, as Verkuyl (1972) first demonstrated. This means that the lexical aspect of a situation can change from one category to another by altering one of the arguments of the verb. Take, for example the following two sentences (Dowty 1986, p. 39):

(3.2)   *John* noticed the rare seashell on the beach (achievement)
          Tourists noticed the rare seashell on the beach (activity)

What this example shows is that an indefinite plural NP, in this case *tourists*, can turn an achievement into an activity.

Despite the fact that this notion of aspect being compositional in nature has been widely accepted (Verkuyl 1989), throughout the second language acquisition studies on aspect researchers often use misleading terminology, referring to the lexical aspect of *verbs* or *predicates* exclusively. Li & Shirai (2000, p. 18) acknowledge this confusion in their own work in the following remark:

> We have been using the term "lexical aspect of verbs", but strictly speaking, the lexical aspect value is determined by both the verb and its arguments, which Smith (1997) calls the verb constellation. Examples include [John love Mary] (state) (...) Note that what is inside [] is not a linguistic form but the proposition underlying it, without any verb morphology to signal the viewpoint aspect. In such instances, we are talking about the semantic structure of verb-plus-arguments without any value imposed by grammatical aspect.

Grammatical aspect is another component in the composition of aspect, a phenomenon to which we now turn.

Smith (1991) likens grammatical aspect to the lens of a camera. Just as a camera lens focuses an object, so grammatical aspect focuses a situation either in part or as a whole. It is encoded grammatically, typically via auxiliaries and/or inflectional and derivational morphology. Types of grammatical aspect, or aspectual viewpoints, include perfective, imperfective and neutral[9] which Smith (1991, p. 6) defines as follows:

> *Perfective* viewpoints focus on the situation as a whole, with initial and final points.
>
> *Imperfective* viewpoints focus on part of a situation, including neither initial nor final point.
>
> *Neutral* viewpoints are flexible, including the initial point of a situation and at least one internal stage (where applicable).

The following two sentences contrast perfective and imperfective aspect in English:

(3.3)   John built a house last spring (perfective)

            John was building a house last spring (imperfective)

The first sentence is viewed externally; that is, as a completed whole. It means that John finished building the house. The second is viewed internally, meaning that while we know that John started building a house last spring, whether or not it was completed is unknown.

To illustrate the fundamental contrast in grammatical aspect between perfectivity and imperfectivity, Comrie (1976, pp. 24-25) proposes the classification of aspectual oppositions displayed in Figure 3.1. This classification accounts for the variety of ways in which imperfective aspect is expressed across languages. It is divided into the categories of habitual and continuous. Habitual situations describe an iterative situation which occurs over an extended period of time. Continuous situations are either progressive or non-progressive, both of which

---

[9]We will not address neutral viewpoint in any detail as it is not addressed in the second language acquisition research, nor is it applicable to past tense aspect in French, the target language of our investigation. We refer the interested reader to Smith (1991, p. 119-123).
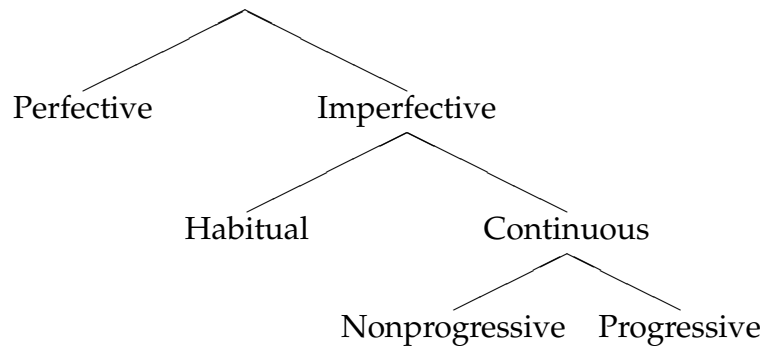
Figure 3.1: Classification of Aspectual Oppositions

express situations in progress without initial or final endpoints. They are distinguished by the fact that progressive aspect is not compatible with states.

Comrie (1976) observes that languages express imperfectivity, in particular, in three main ways. Some languages have a single category of imperfectivity which encompasses the habitual, progressive and non-progressive meanings. Others have a distinct category for each meaning, while a third type has a category that corresponds to a particular meaning. With respect to past tense aspect, French expresses imperfectivity in the first way and English in the third way. In addition to explaining that the perfective denotes a complete situation with a beginning, middle and end, Comrie (1976) points out that in many languages states marked with perfective aspect can have an ingressive meaning (i.e. it indicates the beginning of the state). This is a feature of French, but not English. These differences in the expression of grammatical aspect in French and English will be discussed in detail in the next section.

3.3   The French and English Aspectual Systems

The description of the aspectual systems of English and French will pertain only to aspect as it is expressed in the past tense (for work which discusses other tenses see Jespersen 1931, Joos 1964, Smith 1991, Twaddell 1965). Discussion will be restricted to grammatical aspect, as this is where the two aspectual systems differ most starkly and, therefore, the most likely place for transfer problems to arise. Also for this reason, it will focus on the significant ways in which the two systems are different, rather than comprising an exhaustive comparison. We will

follow the analysis of Smith (1991), as she presents the aspectual systems of both French and English within a single theoretical framework.

3.3.1 Aspect in English

In English, perfective aspect is available for all categories of lexical aspect. It is distinct from other languages in that the aspectual value of the perfective is determined by the category of lexical aspect to which the verb constellation belongs. Imperfective aspect in English typically has a progressive meaning which is unmarked on non-states. We will consider perfective and imperfective aspect separately in the following two sections.

Perfective Aspect

As discussed in Section 3.2.1, the categories of lexical aspect (i.e. states, activities, accomplishments and achievements) are defined by the temporal properties of dynamism (±static), durativity (±durative) and telicity (±telic). Each category of lexical aspect conforms to a particular configuration of temporal properties. For instance, an activity is a durative (+durative) event (-static) with an arbitrary final endpoint (-telic). Following this pattern of temporal properties, the verb constellation [Bill walk in the park] can be categorized as an activity because it takes time, is dynamic and has an arbitrary final point.

According to the definitions presented in Section 3.2.2, grammatical aspect is characterized by the endpoints (initial or final) of a situation. Perfective aspect includes the initial and final points of a situation, while imperfective aspect does not. In English, perfective aspect is formed using the past simple form of the main verb and is available to all categories of lexical aspect, as the following examples show:

(3.4)   a.   Jennifer jogged in the park (activity)
         b.   Fiona walked to school (accomplishment)
         c.   John won a squash game (achievement)
         d.   Mary hated housework (state)

As mentioned at the beginning of Section 3.3.1, perfective aspect is distinctive in English because its aspectual value varies according to each category of lexical

aspect[10]. Specifically, a sentence exhibiting perfective aspect maintains the endpoint properties of the lexical aspectual category of its verb constellation. For instance, the verb constellation [Fiona walk to school] is an accomplishment. As such, it has a natural final endpoint. The aspectual value of the perfective, when applied to an accomplishment (see Example 3.4b above), includes this endpoint property. Similarly, activity sentences with perfective aspect include an arbitrary final endpoint (see Example 3.4a above) and achievement sentences with perfective aspect convey a single-stage event with a natural final endpoint (see Example 3.4c above). Therefore, activities, accomplishments and achievements are considered closed or complete when expressed using perfective aspect.

Since the lexical aspectual category of states does not include endpoints, stative sentences with perfective aspect are considered to have open endpoints. As such, they can have more than one interpretation. Consider the following examples from Smith (1991, p. 221):

(3.5) a. Sam owned three peach orchards
       b. Mary lived in New Orleans
       c. Bill was angry

For each of these sentences, two interpretations are possible. The first interpretation is that the state continues into the present. For Example 3.5a, this would mean that Sam owned three peach orchards at some time in the past and still owns them now. The second interpretation is that the state has ended. Taking the same example, this would mean that Sam owned three peach orchards at some time in the past but no longer owns them now.

Imperfective Aspect

As discussed in Section 3.2.2, imperfective aspect focuses on part of a situation without any information about either initial or final endpoints. In English, im-

---

[10]We note that not all semanticists approach this dimension of aspect in English in the same way. de Swart (1998), for instance, takes tense as her starting point, arguing that the past simple in English applies to states, processes (i.e. activities) and events (i.e. achievements and accomplishments) and is aspectually neutral. In spite of this, she comes to the same conclusion, namely that the aspect of sentences in the past simple is determined by the lexical aspect of the verbal predicate.

perfective aspect is mainly expressed by the progressive[11]. The progressive is formed with the past simple of the auxiliary *be* followed by the main verb ending in *-ing*. More than simply the expression of a situation in progress, progressiveness is defined as the combination of progressive meaning and non-stative meaning (Comrie 1976). As such, it applies only to non-stative sentences, as the following examples show (Smith 1991, p. 113 and p. 223):

(3.6) a. Kelly was singing (activity)
    b. Ross was climbing a tree (accomplishment)
    c. Bright Star was winning the race (achievement)
    d. *Bill was knowing the answer (stative)

The progressive focuses on the internal stages, or duration, of non-stative situations. As discussed in Section 3.2.1, durativity is the distinguishing feature between achievements and accomplishments; accomplishments have duration, while achievements do not. To explain why imperfective aspect is felicitous with achievements, Smith (1991) argues that imperfective aspect focuses on the preliminary stages of achievements. Other researchers (Dowty 1986, Verkuyl 1989) have used the same evidence to claim that achievements do have duration and need not constitute a separate lexical aspectual category.

It is possible to apply progressive aspect to statives; however, such usage is marked. Consider the following examples (Smith 1991, p. 226) (with modifications from present to past tense):

(3.7) a. John was really liking the play
    b. The cake was looking done

These examples are dynamic and have a temporary quality which is characteristic of events rather than states. For instance, Example 3.7a conveys the idea that John was active in liking the play but that this may not necessarily continue.

Habituality is not expressed with imperfective aspect in English. English does have the form *used to* to convey habituality in the past tense (e.g. When I was

---

[11]English also conveys imperfective aspect via the resultative imperfective (e.g. The picture was hanging on the wall (Smith 1991, p. 224)). It will not be discussed here, as it is a specific type of imperfective aspect which did not appear in our data.

small, I used to go to Scotland every summer). However, the past simple, which typically expresses perfective aspect, can also be used (e.g. When I was small, I went to Scotland every summer). This point will be taken up in more detail in Sections 3.3.2 and 3.5.

3.3.2 Aspect in French

The aspectual system of French has been studied at great length (see work by Benveniste (1966), Garey (1957), Grévisse (1949), Guillaume (1929), Hoepelman & Rohrer (1980), Imbs (1960), Sten (1952) and Vet (1980)). It is unique in that perfective aspect presents all situations as closed and imperfective aspect presents all situations as open.

Perfective Aspect

In the past tense, the perfective is expressed via the *passé simple* or the *passé composé* and has the same aspectual value in either tense[12]. For this reason, and since our experiment explicitly excluded the *passé simple*, it will not be discussed further. The *passé composé* is formed using the present tense of the auxiliary verb *être* or *avoir* followed by the past participle of the main verb. All categories of lexical aspect can be expressed using perfective aspect, as the following examples illustrate (Smith 1991, p. 255, modified for consistency with Example 3.9. Note that the superscripts ^{PC} and ^{Pres} will be used to indicate *passé composé* and present tense, respectively):

(3.8)   a.   Ce matin Marie a chanté (activity)
               This morning Marie sang^{PC}

        b.   L'été passé ils ont construit une cabine (accomplishment)
               Last summer they built^{PC} a cabin

        c.   La guerre a éclaté (achievement)
               The war broke out^{PC}

        d.   Jean a été malade hier soir (state)
               Jean was^{PC} sick last night

---

[12]This is accurate, provided that we exclude the meaning of the *passé composé* that considers the completed situation as being related to the present, as in *J'habite en Ecosse depuis 2002*. English expresses this meaning via the present perfect: *I have lived in Scotland since 2002*.

The perfective in French presents all categories of lexical aspect as closed with initial and final endpoints. Smith (1991) shows that this is the case by using a conjunction test. According to this test, perfective sentences cannot be conjoined felicitously with assertions that the situation in question will continue. Consider the following examples (Smith 1991, p. 255):

(3.9) a.   # Ce matin Marie a chanté; peut-être qu'elle chante encore
            This morning Marie sang^{PC}; perhaps she is still singing^{Pres}

       b.   # L'été passé ils ont construit une cabine; peut-être qu'ils la con-
            struisent encore
            Last summer they built a cabin^{PC}; perhaps they are still building^{Pres}
            it

       c.   # La guerre a éclaté; peut-être qu'elle éclate encore
            The war broke out^{PC}; perhaps it is still breaking out^{Pres}

       d.   # Jean a été malade hier soir et il est malade maintenant
            Jean was sick^{PC} last night and he is^{Pres} sick now

Example 3.9a is odd. It proposes that Marie is maybe still singing after having started and finished singing this morning. Examples 3.9b and 3.9c are contradictory. In the case of 3.9b, if they built a cabin last summer, it is complete and, therefore, cannot still be being built now. Similarly, if the war broke out at some moment in the past, the same war cannot break out now, as Example 3.9c suggests.

The oddity of Example 3.9d is not captured by the English translation. In English, *Jean was sick last night and is sick now* is a perfectly natural sentence. The use of perfective aspect with states in French entails a change out of that state. Accordingly, expressing the state [Jean être malade] with perfective past tense aspect means that he was sick and then changed out of the state of being sick. In contrast to the English example, the French cannot mean that Jean is now sick with the same illness that he had earlier.

As mentioned in Section 3.2.2, French has a general imperfective which encompasses habitual, non-progressive and progressive meanings. The general imperfective is expressed via a tense called the *imparfait*. The *imparfait* is formed using a main verb stem followed by a series of affixes. The affixes vary according to person and number with the exception of the first and second person singular which are identical phonetically and orthographically. The imperfective presents situations as continuing, with neither initial nor final endpoints. It is available for activities, accomplishments, achievements and states. We repeat the situations from the examples of imperfective aspect in English (see Example 3.6):

(3.10) a. Kelly chantait (activity)

Kelly was singing^{Impf}

b. Ross grimpait dans un arbre (accomplishment)

Ross was climbing^{Impf} a tree

c. Bright Star gagnait la course (achievement)

Bright Star was winning^{Impf} the race

d. Bill savait la réponse (state)

Bill knew^{Impf} the answer

While in French the *imparfait* appears on all four of these examples, the progressive has been used in the English translation for the activity, accomplishment and achievement, but not the state. This highlights an important difference between French and English which will be taken up in Section 3.5. As mentioned in Section 3.2.2, progressive and non-progressive meanings of the imperfective are distinguished by the fact that the former is compatible only with non-states, while the latter is compatible only with states. Therefore, in the *imparfait*, the state in example 3.10d expresses non-progressive aspectual meaning, while the other examples (3.10a, b and c) convey progressive aspectual meaning.

For all of the examples in 3.10, there is no indication of when the situation started or whether it might have ended. In other words, it is open without initial nor final endpoints. This is verified by, again, applying the conjunction test, explained in Section 3.3.2 under *Perfective Aspect* and applied to the perfective in Example 3.9. We will use it to test the imperfective examples in 3.10:

(3.11) a.   Kelly chantait; peut-être qu'elle chante encore

Kelly was singing^{Impf}; perhaps she is still singing^{Pres}

b.   Ross grimpait dans un arbre; peut-être qu'il y grimpe encore

Ross was climbing^{Impf} a tree; perhaps he is still climbing^{Pres} it

c.   Bright Star gagnait la course; peut-être qu'il la gagne encore

Bright Start was winning^{Impf} the race; perhaps he is still winning^{Pres} it

d.   Bill savait la réponse; peut-être qu'il la sait encore

Bill knew^{Impf} the answer; perhaps he still knows^{Pres} it

The fact that all of these examples are felicitous with a present tense conjunction shows that the *imparfait* does not include a final endpoint, regardless of lexical aspectual class.

The absence of initial point is tested by conjoining a sentence in the *imparfait* and a perfective subordinate clause introduced by when (or *quand* in French). If the main clause includes a final point (i.e. is *not* imperfective), the perfective subordinate clause should evoke a sequential interpretation. Consider the following example:

(3.12) Quand l'ours est venu, Ross a grimpé dans un arbre

When the bear came^{PC}, Ross climbed^{PC} a tree

In this case, the main clause, *Ross a grimpé dans un arbre*, includes a final point and it is conjoined to a perfective subordinate clause, *Quand l'ours est venu*. As predicted, the result is a sequential interpretation: first the bear came and then Ross climbed a tree. In other words, the arrival of the bear signalled the initial endpoint of Ross climbing a tree.

Taking the same example, a sequential interpretation will not result, if we substitute a main clause that does not include an initial endpoint:

(3.13) Quand l'ours est venu, Ross grimpait dans un arbre

When the bear came^{PC}, Ross was climbing^{Impf} a tree

Contrary to Example 3.12, Example 3.13 does not mean that Ross began to climb the tree at the moment the bear came. Rather, it means that his climbing was in progress when the bear arrived on the scene.

The French imperfective also expresses progressiveness. As mentioned in our discussion of imperfectivity in English, progressiveness is compatible only with non-stative verb constellations. Since there is a direct correspondence of form and meaning between French and English in the expression of progressiveness, anglophones tend not have difficulty with this particular aspectual meaning (Darbelnet 1977).

French has a further way of expressing progressiveness, that is with the lexical form *être en train de*. It too can only combine with non-stative verb constellations and is considered optional, since the *imparfait* can express the same meaning[13]. The following examples show that progressive meaning in French is only compatible with non-statives:

(3.14)   a.   Kelly était en train de chanter (activity)
              Kelly was in the process of singing
         b.   Ross était en train de grimper dans un arbre (accomplishment)
              Ross was in the process of climbing a tree
         c.   Bright Star était en train de gagner la course (achievement)
              Bright Star was in the process of winning the race
         d.   *Bill était en train de savoir la réponse (state)
              Bill was in the process of knowing the answer

The activity, accomplishment and achievement examples in 3.14 could be expressed in the *imparfait* and maintain the same meaning. The state, however, would only allow a non-progressive interpretation.

The third possible meaning of the imperfective in French is habituality. Here are some examples:

---

[13]Smith (1991) claims that progressive aspectual meaning in French is expressed exclusively through the lexical form *être en train de* (to be in the process or midst of). Her justification is that it conforms to the progressive features of being incompatible with states and conveying a temporary state of affairs. However, her explanation does not account for the fact that the *imparfait* clearly has a progressive meaning in some contexts, e.g. *La directrice défendait sa décision, mais le journaliste l'a interrompue avant qu'elle ait pu terminer* (The director was defending her decision, but the journalist interrupted her before she could finish).

(3.15) a. Lucie nageait dans le lac pendant une heure
            Lucie used to swim^{Impf} in the lake for an hour

        b. Pierre marchait au travail jusqu'à sa retraite
            Pierre used to walk^{Impf} to work until his retirement

        c. Jacques gagnait le championat de fléchettes quand il était jeune
            Jacques used to win^{Impf} the darts championship when he was young

        d. Autrefois, Sylvie était malade pendant des semaines
            In the past, Sylvie used to be^{Impf} ill for weeks

All of these sentences express situations that occurred several times in the past. Although the imperfective has been translated with the expression *used to*, this is an approximation because English does not have an equivalent imperfective form. Depending on the context, other translations for 3.15a, for instance, include *Lucie would swim in the lake for an hour* and *Lucie swam in the lake for an hour (every day)*. As this last translation shows, habitual meaning in English cannot necessarily be conveyed without the help of adverbials or context. This point will be taken up in detail in Section 3.5.

3.4   Aspectual Coercion

Discussion of the French and English aspectual systems in Section 3.3 focused on the unmarked usage of grammatical aspect. We devote this section to exploring the phenomenon of aspectual coercion, including how it relates to grammatical aspect. We will begin by looking at shifts in meaning which occur by syntactic means. We will then go on to define and illustrate aspectual coercion generally and explore particular manifestations of it in French and English.

3.4.1   Shifts in Meaning via Syntax

In Section 3.2.1, we briefly mentioned that aspectual meaning can shift depending on the combination of syntactic constituents present in the utterance. Specifically, we pointed out that changing one of the arguments of the verb from a definite singular NP (John) to an indefinite plural NP (tourists) would shift the category of the sentence from an achievement to an activity. The example is repeated below (Dowty 1986, p. 39):

(3.16)   *John* noticed the rare seashell on the beach (achievement)

            Tourists noticed the rare seashell on the beach (activity)

Several other syntactic constituents which have been identified as contributing to determining aspect are described below with examples taken from the literature:

> Prepositional phrases can turn an activity into an accomplishment, as the following examples illustrate (Dowty 1986, p. 39):
>
> (3.17)   John walked (activity)
>
>   John walked *to the station* (accomplishment)
>
> A quantified NP can turn an activity into an accomplishment (Verkuyl 1993, p. 47):
>
> (3.18)   She ate *sandwiches* (activity)
>
>   She ate *a sandwich* (accomplishment)
>
> The adverbials *for an hour* and *in an hour* differentiate between states/activities and accomplishments/achievements (Dowty 1986, p. 39) :
>
> (3.19)   John read a book *in two hours* (accomplishment)
>
>   John read a book *for two hours* (activity)

3.4.2   Defining Aspectual Coercion

Shifts in meaning can also occur by pragmatic means. Specifically, a contextual reinterpretation process can occur when there is a conflict between the lexical aspectual class of the verb constellation and the aspectual constraints of some other element of the sentence. This process is known as aspectual coercion (de Swart 1998, Moens 1987, Moens & Steedman 1988). Consider the following example (de Swart 1998, p. 359):

(3.20)   John played the sonata for about eight hours

As illustrated in Example 3.19, the duration adverbial *for an hour* is sensitive to the lexical class of the verb constellation with which it combines: it is usually felicitous with states and activities, but not accomplishments and achievements. In Example 3.20, *John played the sonata* is an accomplishment but has been combined with the duration adverbial *for about eight hours*. Pragmatic knowledge suggests

that sonatas do not take eight hours to play once through. Therefore, the achievement of John playing the sonata is coerced into being an activity by giving the sentence an iterative interpretation. In other words, the coerced meaning is that John played the sonata several times in the course of about eight hours.

Although reflexive clitics and the animacy of subject NPs have recently been proposed as means for invoking aspectual coercion (see Slabakova & Montrul (2005) for examples of these in Spanish), we will examine the more robustly documented use of grammatical aspect and adverbials to evoke aspectual coercion in more detail, contrasting French and English.

3.4.3   Aspectual Coercion: French versus English

States to Inchoative Achievements

States can be coerced into achievements in both French and English; however, each language does so in a different manner. Consider the following example (Smith 1991, p. 255):

(3.21)   Paul a été^{PC} fâché quand Jeanne a cassé l'assiette
          Paul was angry when Jeanne broke the plate

The verb constellation [Paul be angry/Paul être fâché] *per se* is a state. If we considered [Paul be angry] to be a state in the context of Example 3.21, he would already have been angry for some unrelated reason before Jeanne broke the plate. This is possible, but in the absence of further context the more likely interpretation is that Paul was angry because Jeanne broke the plate. Thus, what would normally be considered a state is interpreted as the initial point of an achievement in this example. English achieves this meaning by relying purely on context. French, on the other hand, leaves no room for ambiguity by explicitly marking the state with perfective aspect.

In addition to context, English uses adverbials as a means to the same end, while French relies on perfective aspect. Consider the following example (de Swart 1998, p. 359):

(3.22)   Suddenly, I knew the answer

[I know the answer] is clearly a state. However, the addition of the adverb *suddenly* conveys the idea that there was a moment when the person came to realize that he or she knew the answer. In other words, the state of knowing the answer has been coerced into the initial point of an achievement. French expresses this meaning using perfective aspect without needing the adverb:

(3.23)   (Tout d'un coup) j'ai su^{PC} la réponse

Accomplishments to Activities

As explained in Example 3.20, the conflict between an accomplishment, such as *John played the sonata* and the duration adverbial *for + X amount of time*, which is normally felicitous with states and activities, can be resolved by reinterpreting the sentence as an activity. The same applies in French:

(3.24)   John a joué^{PC} la sonate pendant à peu près huit heures
          John played the sonata for about eight hours

Like the English, the French sentence is coerced from an accomplishment into an activity where John's playing of the sonata is iterative. As discussed in Sections 3.3.2 and 3.5, while English conveys habituality in several ways, it is typically only expressed with imperfective aspect in French. It is possible for Example 3.24 to appear with imperfective aspect:

(3.25)   John jouait^{Impf} la sonate pendant à peu près huits heures
          John used to play/played/would play the sonata for about eight hours

The difference in meaning between the perfective version and this version is that the former is bounded, meaning that it is no longer the case that John is playing the sonata now, whereas the latter, is unbounded, meaning that he may still be playing the sonata. The English version is ambiguous in this respect. Note that only Example 3.24, is an instance of aspectual coercion.

Accomplishments can also be coerced into non-iterative activities in both English and French. Consider the following example (de Swart 1998, p. 359)[14]:

---

[14]We note that this and Example 3.27 are only presented in English in the article by de Swart (1998), even though they are examples of coercion in both French and English.

(3.26)   J'ai lu^{PC} un livre pendant quelques minutes
            I read a book for a few minutes

The same factors are involved in this sentence being an example of aspectual coercion: there is a conflict between the accomplishment, *I read a book/J'ai lu un livre*, and the duration adverbial, *for a few minutes/pendant quelques minutes*. Contrary to the sonata sentence (see Examples 3.20 and 3.24), pragmatic knowledge indicates that usually it takes more than a few minutes to read a book. Thus, the accomplishment is coerced into a non-iterative activity in which the person engaged in reading part of a book. Like for Example 3.24, imperfective aspectual marking on the French version of the sentence is possible and would result in an unbounded interpretation, but would not be an instance of aspectual coercion.

Achievements to Activities

Finally, achievements can be coerced into activities as the following example illustrates (de Swart 1998, p. 359):

(3.27)   Pendant des mois, le train est arrivé en retard
            For months, the train arrived late

Since [the train arrive late/le train arriver en retard] is an instantaneous event, it is considered to be an achievement, rather than an accomplishment (we refer the reader back to Section 3.2.1 for definitions of the Vendlerian categories of lexical aspect). It is incompatible with the duration adverbial, *for months/pendant des mois*. Knowing that *the train arrived late/le train est arrivé en retard* can only refer to a single event, we come to an interpretation of the sentence in Example 3.27 by taking it to mean that a particular train regularly scheduled to arrive at a certain time, arrived late over a period of months.

3.5   Aspect and the Second Language Learner

The fundamental difficulty of aspect, from a transfer perspective, is that the past simple can be equivalent either to a *passé composé* or an *imparfait* (Darbelnet 1977). In other words, perfective forms in English can have perfective or imperfective meaning. In French, on the other hand, there is a clear and consistent relationship

between aspectual form and aspectual meaning: perfective forms only express perfective meaning and imperfective forms only express imperfective meaning. Having explored the aspectual systems of French and English, we can identify the specific contexts in which problems due to language transfer may arise.

3.5.1 Non-Progressive Aspectual Meaning

As explained in Sections 3.3.1 and 3.3.2, English uses verbal morphology to distinguish between progressive and non-progressive imperfective meaning, while French has one form for both meanings. Consider the sentences in French and English below, reproduced from Examples 3.10 a-d:

(3.28)   a.   Kelly chantait (activity)
             Kelly was singing^{Impf}

       b.   Ross grimpait dans un arbre (accomplishment)
             Ross was climbing^{Impf} a tree

       c.   Bright Star gagnait la course (achievement)
             Bright Star was winning^{Impf} the race

       d.   Bill savait la réponse (state)
             Bill knew^{Impf} the answer

The sentences expressing progressive meaning, namely 3.28a-c, have imperfective aspectual marking in both languages: English uses the progressive and French the *imparfait*. To express non-progressive meaning, however, English relies on lexical aspect but uses the past simple which occurs with other categories of lexical aspect to convey perfective aspect. French, on the other hand, maintains the *imparfait*, consistent with the rest of the examples. This divergence in the expression of non-progressive aspectual meaning may be problematic for anglophone learners. Instead of extending the use of the *imparfait* to states in French, they may use the *passé composé* as an equivalent of the past simple. Proponents of the aspect hypothesis may take issue with this claim since some have observed the tendency for learners to overmark states with imperfective grammatical aspect (e.g. Salaberry (1998)). However, in her review of the aspect hypothesis studies on imperfective marking, Bardovi-Harlig (2000, p. 236) states that "tokens of states in interlanguage are typically dominated by *be* and *have* and their equivalents" and "the number of different stative verbs is often limited to about half a

dozen". Thus, it could be that overmarking of states with the French imperfective is limited to specific high frequency verbs and that otherwise learners will use the *passé composé*.

3.5.2 Habituality

The habitual meaning of the *imparfait* could be another challenging dimension of French aspect for anglophone learners. As mentioned in Section 3.3.2, although English has the form *used to* to express habituality in the past tense, the perfective can also be used. Since English also uses the perfective to convey closed situations, the habitual interpretation often has to be recovered from context or adverbials. French, on the other hand, uses perfective aspect to convey closed situations and imperfective aspect to express habitual meaning. Consider the following example (Darbelnet 1977, p. 126):

(3.29)   When the room was warm, the maid brought in breakfast

In English, this could refer to a closed event or to a habitual action, depending on the context. In French, the perfective meaning of a closed event can only be expressed with perfective aspect:

(3.30)   Quand la pièce a été bien chaude, la bonne a apporté le petit déjeuner
          When the room was^{PC} warm, the maid brought in^{PC} breakfast

Likewise, the habitual meaning can only be conveyed with imperfective aspect:

(3.31)   Quand la pièce était bien chaude, la bonne apportait le petit déjeuner
          When the room was^{Impf} warm, the maid brought in^{Impf} breakfast

Anglophone learners may transfer the habitual meaning of the perfective from English and attempt to use perfective aspect to express habituality in French.

3.5.3 Aspectual Coercion

In our comparison of aspectual coercion in French and English in Section 3.4.3, we found that the two languages differed only when coercing states into inchoative achievements. To achieve this type of coercion, French uses grammatical

aspect while English relies on either context (see Example 3.21) or adverbs (see Example 3.22). Because of this difference, transferring the English way of achieving coercion is a possibility that may prove problematic for anglophone learners.

3.6 Chapter Summary

Despite the fact that the difficulty of acquiring aspect in a second language is well-known, little work has attempted to explain why this might be the case. Notably, a single article, that of Andrews (1992), has addressed the question with respect to anglophones learning aspect in French. Thus, taking language transfer as a possible source of difficulty, the goal of this chapter has been to compare the aspectual systems of French and English with a view to establishing what particular dimensions of aspect might be problematic. Our analysis predicted that anglophone learners may have difficulty in comprehending or producing grammatical aspect in French in the following areas:

- use of perfective aspect to mark non-progressive meaning
- use of perfective aspect to mark habituality
- aspectual coercion of states into inchoative achievements

These three manifestations of aspect will serve as the target structures for the experiment described in Chapter 5.

CHAPTER 4

Aspect in Second Language Acquisition

4.1 Introduction

Research on the tense-aspect system in second language acquisition over the past 25 years has generated a wealth of literature, both descriptive and theory-driven[1]. The descriptive work, conducted by the researchers on the European Science Foundation (ESF) Project, has described what linguistic devices second language learners use to express temporality (see Section 4.2). The theoretically-motivated portion of the literature has tested the predictions of the aspect hypothesis (Section 4.3), the discourse hypothesis (Section 4.4) and generative linguistic theory (Section 4.5). Reviewing this research will allow us to identify what further work is needed generally and to motivate research which investigates the effect of feedback on the acquisition of second language aspect (see Section 4.6).

4.2 Descriptive Work: The ESF Studies

An influential body of research on temporality resulted from a large-scale ESF project on second language acquisition by adult foreign immigrant workers in five industrialized Western European countries. Rather than testing a particular theory or focusing only on the structural characteristics of learner expression, the

---

[1]Prior to that, research was focused on the accuracy with which learners produced specific morphemes, rather than on verbal morphology as the manifestation of an underlying semantic system (Bardovi-Harlig 2000).

researchers took a discourse-based functional approach to the study of temporality. Named the "conceptual approach"[2], the idea is to observe and explain how learners express a conceptual domain, in this case temporality, at various points in time.

The data used for the studies were cross-linguistic and longitudinal. The target languages were English, German, Dutch, French and Swedish. In keeping with the goals of the overall project, they were chosen because they encompassed the languages learned by the largest populations of immigrants. The immigrants themselves were young (ages 17-38) working-class monolinguals who came from a range of first language backgrounds, including Punjabi, Italian, Turkish, Arabic, Spanish and Finnish. Living in the host country, they were learning the target language naturalistically and were minimally proficient at the beginning of the project.

The studies were organized so that speakers of two typologically different languages could be observed learning a common second language which was more closely related to one of the source languages than the other. For instance, native speakers of Spanish and Arabic were observed learning French as a second language, to which Spanish is more closely related than Arabic. The learners were recorded engaging in real-life, day to day exchanges or in equivalent role plays with researchers and volunteers, such as bank clerks and employment officers. They also recounted personal experiences and retold films[3]. A total of 20-25 recordings approximately two hours in length were collected from forty learners over a period of two and a half years.

Dietrich et al. (1995) proposed three major stages in the acquisitional process of temporality: (a) pre-basic varieties, (b) the basic variety and (c) further development. Learners of the pre-basic and basic varieties do not exhibit tense or aspect marking. Since our learners were beyond these stages and the experiment focused on the effect of correction on aspect as it is encoded verbally, our review will concentrate on the further development stage.

---

[2]This approach is known by several names. See Bardovi-Harlig (2000, p. 22).

[3]The silent films *Modern Times* by Charlie Chaplin, *The Pear Story*, *The Sorcerer's Apprentice*, *The Tin Toy*, *The Thief of Baghdad* and *The Pink Panther* have been used for this purpose in many studies (Bardovi-Harlig 2000), regardless of their general approach to the study of aspect in second language acquisition.

4.2.1 The Pre-Basic Variety

The pre-basic variety is characterized by the use of uninflected nouns, adjectives, verbs, adverbials and a few particles, such as those for negation. At this stage, learner language exhibits no syntactic structure. Therefore, they must use adverbial expressions, such as days of the week, specific years, time of day, etc. and pragmatic means of expression, specifically the principles of discourse organization. Dietrich et al. (1995) highlight the principle of natural order, according to which the sequence of events are reported in the order in which they happened. In example 4.1, we arrive at the interpretation "after I finished school, I came to Germany" by following the principle of natural order (von Stutterheim & Klein 1987, p. 198):

(4.1)   Schule fertig, Deutschland komm.
         "school finish, Germany come."

Meisel (1987) reports two further principles of discourse organization used by learners in the very early stages of acquisition: scaffolded discourse and implicit reference. In scaffolded discourse, learners rely on the contributions of their interlocutors to express temporality. Typically, this means that when the learner does not understand a question or is having difficulty forming an utterance, the interlocutor will suggest possible answers or interpretations for the learner to choose from. When using implicit reference, the learner uses shared knowledge or context to enable the interlocutor to situate the event appropriately on a timescale. For instance, in the following example, the learner establishes the context as being when his sister was small, thereby signalling to the interlocutor that he is referring to the past (Meisel 1987, p. 213):

(4.2)   meine schwester klein meine mutter ni arbeit
         "my sister small my mother not work."

4.2.2 The Basic Variety

At the basic variety stage, uninflected verbs appear with their arguments in a default form. Depending on the target language, possible candidates for the default form are the bare stem (English, Dutch), V-*ing* (English), infinitive (German,

French, Dutch) or a generalized inflected form (Swedish). At this stage, more adverbials (e.g. *before*, *after*) and connectives (e.g. *and then*) begin to appear, as do boundary markers, such as *start* or *finish* which allow marking of the beginning or end of a situation (e.g. *work finish*, "after working is/was/will be over") (Klein et al. 1995).

4.2.3   Further Development

It is beyond the basic variety stage that learners begin to mark verbal morphology. Klein et al. (1995) make four observations about the acquisition of temporality at this stage. Firstly, progress is slow, gradual and continuous. This is the case for vocabulary, such as adverbials, and for producing verbal morphology. They note that learners go through long periods of co-occurring correct and incorrect usage.

Secondly, mastery of verb forms often precedes mastery of their function. For instance, learners of English will use a bare stem, a V-*ing* form or a variety of present tense forms indiscriminately, without any clear indication that they acknowledge the different function of each. In the following excerpt of an example given by Klein (1995), Madan, a Punjabi-speaking learner of English, recounts part of a Charlie Chaplin silent film in which Charlie is trying to get himself arrested. Madan uses either a base form or a V--*ing* form without any apparent difference in temporal or aspectual function (Klein 1995, p. 60) (numbering altered from original):

(4.3)   (1) next door + shop cigar + cigarettes + everything buy
           (2) charlie say 'shopkeeper + give me one cigar'
           (3) he give it
           (4) he smoking
           (5) after + two baby < = small children > coming
           (6) hes coming
           (7) charlie say 'what you want'
           (8) charlie pickerup cigar/cigarette
           (9) 'here you are and here you are'
           (10) 'right + you go on'
           (11) children go

(12) charlie say 'how much money?'

(13) shopkeeper say 'like this'

(14) charlie say policeman 'give the/this money'

(15) police station police car coming

(16) policeman say 'go on + stay in the car' [...]

(17) charlie stay in the car

(18) stand and there inside stan/sit down fat woman $<$ = he sat down on her $>$

(19) he $<$ = she $>$ pushing charlie

(20) charlie go other side

Overall, the base forms (seen in (2), (3), (7), (8), (11), (12), (13), (14), (16), (17), (18) and (20)) mark the events of the plot as it unfolds, which is often achieved by citing direct speech. Close inspection of the V-*ing* forms (seen in (4), (5), (6), (15), (19)) indicates that they too mark the next event in the story. In (4), for example, it seems that Madan is trying to express the second of two consecutive events: the first, that the shopkeeper gave Charlie a cigar and the second, that he smoked it. Similarly, (5) is the next event in the sequence and is clearly marked as such by the word *after* (i.e. after that/then). (6) is rather difficult to interpret, but appears to be a repetition of (5) with a kind of conflated pronoun *hes* which is apparently common in this learner's discourse and can mean *he is, he was, they have, he has*. (15) is another event in the plot. (19) is the second in a series of three connected events. Firstly, Charlie sat down on the fat woman, then she pushed him off, then Charlie went to the other side. In summary, it appears that Madan is using two forms without a distinguishing function. This observation has been quantitatively corroborated by Bardovi-Harlig & Bofman (1989) who found 7.5 times more errors in the use of tense-aspect morphology, as compared to the form, following an error analysis of composition exams written by advanced learners of English. Furthermore, in a study on tense-aspect use, Bardovi-Harlig (1992a) reported a significant difference in the production of accurate forms as compared to their appropriate use.

The third observation made by Dietrich et al. (1995) about the further development stage is that irregular morphology precedes regular morphology. Two main explanations for this have been proposed. It could be that irregular verbs

are more frequent in the input (Dietrich et al. 1995) or that they are phonologically and perceptually more salient than their regular counterparts (Bayley 1991, 1994, Dietrich et al. 1995, Wolfram 1985, 1989). For the learners in this particular study, there is also the possibility that when learning a second language naturalistically, it is easier to learn individual irregular items, rather than to induce a rule about how to form the regular tense from the input.

Fourthly, Dietrich et al. (1995) observe that learners mark tense using morphological means before they mark aspect. This claim is in contrast to that of the defective tense hypothesis (Andersen 1986, 1991) and the first articulation of the primacy of aspect hypothesis (Robison 1990), the precursors to the aspect hypothesis.

4.3   The Aspect Hypothesis Studies

The aspect hypothesis grew out of work in first language acquisition and in creoles (see Andersen (1989) and Andersen & Shirai (1996) for reviews of work in both fields) and has undergone several revisions since its conception. In early work, it was referred to as the defective tense hypothesis (Andersen 1986, 1991), following Weist et al. (1984) who proposed a similar hypothesis by the same name for first language acquisition. The claim of the defective tense hypothesis in second language acquisition was that "in beginning stages of language acquisition only *inherent aspectual* distinctions are encoded by verbal morphology, not tense or grammatical aspect" ((Andersen 1991, p. 307), emphasis in the original). In the same vein, Robison (1990, p. 316) proposed the primacy of aspect hypothesis which states that "target language verbal morphemes, independent of their function in the target language, are first used by the learner to mark aspect". This prediction was subsequently revised as a result of the findings of the (1990) study to say that "verbal morphology correlates with lexical aspect at least during some stage in the development of IL" (Robison 1990, p. 330). Subsequent research incited further revision, resulting in the current formulation, termed the aspect hypothesis. It is articulated as follows (Andersen & Shirai 1994, p. 133):

> First and second language learners will initially be influenced by the inherent semantic aspect of verbs or predicates in the acquisition of tense and aspect markers associated with or/affixed to these verbs.

More specifically, the hypothesis predicts perfective marking with events, imperfective marking with states and progressive marking with activities as the following four claims of the hypothesis (Shirai 1991, pp. 9-10 as cited by Bardovi-Harlig 2000, p. 227) express[4]:

1. Learners first use (perfective) past marking on achievements and accomplishments, eventually extending use to activities and statives.

2. In languages that encode the perfective/imperfective distinction, imperfective past appears later than perfective past, and imperfect past marking begins with statives, extending next to activities, then to accomplishments, and finally to achievements.

3. In languages that have progressive aspect, progressive marking begins with activities, then extends to accomplishments and achievements.

4. Progressive markings are not incorrectly overextended to statives.

4.3.1   Methodology of the Aspect Hypothesis Studies

Testing of the aspect hypothesis has included empirical work on the target languages of English, Spanish, French, Italian, Japanese and, to a lesser extent, Portuguese, Catalan and Dutch. Both instructed and uninstructed learners have been used. Typically, the untutored learners were the subject of early case studies (e.g. Andersen 1986, 1991 on children and Flashner 1989, Kumpf 1984 on adults) while instructed learners were recruited for the large-scale cross-sectional studies (e.g. Bardovi-Harlig 1992a, Bardovi-Harlig & Bergström 1996, Bardovi-Harlig & Reynolds 1995, Bergström 1995, 1997, Collins 1997, Hasbún 1995). While there has been some longitudinal work, typically in the form of case studies, most experiments have been cross-sectional in design, taking advantage of large numbers of learners at different levels within a university setting. Data has been

---

[4]These four claims are also cited by Andersen (2002, p. 79). That version was not reproduced here because it was expressed less succinctly, owing to the fact that it was edited slightly from the version applicable to first language acquisition (Andersen & Shirai 1996, p. 533).

elicited according to a number of procedures, including usually some combination of oral and written, personal and impersonal narratives, written cloze passages[5] and judgment tasks.

A study testing the aspect hypothesis consists of several stages. First of all, the level of the learners is determined according to, for example, course level or the results of a grammar test. Data are then elicited using one or more of the techniques described above. If the task allows relative freedom of expression on the part of the learner, as in narratives for example, each sentence must be coded according to its aspectual class. For other tasks, such as written cloze passages, the aspectual class of each sentence will be pre-determined by the researcher. Once the data are collected, they are coded according to the lexical aspectual class of each sentence, using, in most cases, diagnostic tests such as those described in section 3.2.1. Most studies use the Vendlerian categories of lexical aspect; however, there are some which opt for binary classifications (Bayley 1994, Giacalone Ramat & Banfi 1990, Kaplan 1987 telic vs atelic; Robison 1990 punctual vs durative), three-way classifications (Ramsay 1990, Salaberry 1999 events/telics, activities, states) and six-way classifications (Giacalone Ramat 1995, 1997, Leiria 1994, Leiria & Mendes 1995). After the data are coded for lexical aspect, they undergo morphological coding which will identify the category of grammatical aspect of each sentence. Finally, the analysis examines the distribution of verbal morphology within and across lexical aspectual categories. Results which support the aspect hypothesis will show, for instance, a progression of perfective marking from achievements to accomplishments to activities that is positively correlated with learner proficiency (i.e. the more proficient the learner, the further the progression).

4.3.2 Findings of the Aspect Hypothesis Studies

Perfective marking has been the most widely studied form of grammatical aspect in the aspect hypothesis studies. Bardovi-Harlig (2000) suggests that this may be the case because perfective past marking appears first and because it is the most common type of marking in narratives, on which most samples are based.

---

[5]The cloze passages were also narratives.

Bardovi-Harlig & Reynolds (1995) used cloze passages to elicit data from ESL learners at six different placement levels and with a variety of first languages. They found that learners at all levels used the past simple more with telic than atelic verbs but that this gap narrowed as proficiency level increased. Collins (1997) tested seventy francophone learners of English. One group underwent the same procedure as the participants in the Bardovi-Harlig & Reynolds (1995) study while the other performed basically the same task but with cloze passages that had an equal representation of aspectual classes (9 tokens from each class) as well as a wider variety of distracters. She found that learners used the simple past significantly more frequently with events as opposed to activities or states.

Bardovi-Harlig (1998)'s study analyzed both written and oral narratives. While the results of both tasks showed a progression of the past simple from achievements to accomplishments to activities, her analysis of the written narratives found that achievement and accomplishment predicates tended to pattern together as events, whereas the oral narratives exhibited a clearer gradation between aspectual categories. Using data exclusively from oral interviews produced by twenty-six Puerto Rican learners of English at four university levels of proficiency, Robison (1995) found that of all the lexical aspectual categories past simple marking occurred most frequently on event predicates.

In their study of classroom learners of English and French, Bardovi-Harlig & Bergström (1996) found that on a written story retell, learners of both languages at the lowest level showed a concentration of perfective marking (i.e. past simple/passé composé) on achievements, learners of English tended to mark the progressive on activities and learners of French favoured imperfective inflections on states. In addition, as their level increased, learners' perfective marking spread from achievements to accomplishments, but progressive and imperfective inflections were resistant to extension to other lexical aspectual categories. It was also the case that progressive forms (in English) were the second most common marking on activities, next to base forms, and that only students at the highest level showed use of the progressive with accomplishments.

Bergström (1995) collected written narratives and cloze passages based on *Modern Times*, a silent film by Charlie Chaplin, from 117 anglophone learners of French at first, second and third year university levels. In the narrative task,

she found a strong tendency to use the perfective with achievement, accomplishment and activity verbs already among the first year learners. However, results from the cloze task, which included a broader sampling of predicates than the narrative task, showed that when second and third year learners produced the passé composé, they clearly distinguished achievements from accomplishments and activities, supporting the aspect hypothesis.

Kihlstedt (2002) looked at data from four interviews over two years in which four advanced learners of French and four native speakers discussed their personal history, general opinions of Sweden and France, previous academic and professional experience, hobbies and future plans. She found that perfective marking by learners conformed to the predictions of the aspect hypothesis whereas native speakers used perfective marking to roughly the same degree to mark both statives and activities.

Support for the aspect hypothesis about the spread of the imperfective from statives to activities to accomplishments to achievements also exists. Hasbún (1995) collected written narratives from eighty anglophone learners of Spanish at first, second, third and fourth year university levels as well as from a control group of native speakers. She found that imperfective marking of states emerged in the third year and spread to activities in the fourth year. Similarly, in the written narratives from her study described above, Bergström (1995) observed imperfective marking on states in the second year learners which spread to activities in the third year. Wiberg (1996)'s study of Italian learners of Swedish-Italian heritage living in Sweden found that the learners strongly associated the imperfective with statives and that one learner in the advanced group of twenty-four subjects used the imperfective with activities. As for the ultimate stage of imperfective marking on achievements, Cadierno (2000) found the use of the imperfect spreading to include achievements among a group of advanced Danish learners of Spanish.

A few studies have investigated the development of the progressive. Giacalone Ramat (1997)'s cross-sectional study of twenty learners of Italian of mixed first language backgrounds involved a conversational interview based on a film retell and a description of picture stories. Consistent with the predictions of the aspect

hypothesis, she found that 63% of all progressive tokens occurred with activities and that progressive marking spread slowly to accomplishments (8%) and achievements (4%). Shirai & Kurono (1998)'s study comprised seventeen tutored Chinese learners of Japanese who performed a judgment task three months after they had arrived in Japan and then twice again at three month intervals. Their findings suggested that learners correctly recognized the progressive marker more often with activities than with achievements.

As the above findings show, support for the aspect hypothesis has been attested in the acquisition of English (Bardovi-Harlig 1998, Bardovi-Harlig & Bergström 1996, Bardovi-Harlig & Reynolds 1995, Collins 1997, Robison 1995), French (Bergström 1995, Bardovi-Harlig & Bergström 1996, Kihlstedt 2002), Spanish (Cadierno 2000, Hasbún 1995), Italian (Giacalone Ramat 1997, Wiberg 1996) and Japanese (Shirai & Kurono 1998) as a second language[6]. In comparison, studies that have found evidence against the aspect hypothesis are scarce. In her monograph on the acquisition of tense and aspect by second language learners, Bardovi-Harlig (2000) discusses work by Kumpf (1984), Schumann (1987), Robison (1995), Rohde (1996) and Salaberry (1999) which pose a potential challenge to the aspect hypothesis. Of these, only Salaberry (1999)'s study on second language Spanish poses a problem for the aspect hypothesis. While overall the study supported the aspect hypothesis, the data produced by learners at the lowest level did not. Use of perfective grammatical aspect occurred across aspectual classes which is not expected for learners at this level. Thus, Salaberry (1999) partially rejected the aspect hypothesis, concluding that learners at the beginning stages of acquisition use verbal morphology independently of the inherent lexical aspect of verbal predicates and suggested that the preterite was being used as a default past tense marker, rather than to express perfective grammatical aspect. Subsequent work by Salaberry (2003) has found further evidence of default past tense marking, although more research is needed on this question.

---

[6]Support has also been found in second language Dutch (Housen 1993, 1994) and Catalan (Comajoan 1998). However, as they were case studies, it would be premature to make any general claims about learners of these languages.

4.4 The Discourse Hypothesis Studies

The discourse hypothesis is a first step towards determining the role of discourse in the acquisition of second language tense and aspect. It is centred on narratives and posits that narrative structure has an influence on aspectual choices. Narrative structure consists of two parts: the foreground and the background. The main events of the discourse which represent new information occur in the foreground, while supporting or elaborative information (e.g. scene setting) constitutes the background (Bardovi-Harlig 1998, 2000). Inspired by work by Hopper (1979), who observed that native speakers will use aspect to mark narrative structure, the interlanguage discourse hypothesis claims that "learners use emerging verbal morphology to distinguish foreground from background in narratives" (Bardovi-Harlig 1994, p. 43). Specifically, verbs in the foreground will carry simple past morphology[7], whereas verbs in the background will not[8] (Bardovi-Harlig 1998, 2000).

Research on the discourse hypothesis has been a mixture of case studies (of second language English (Flashner 1989, Kumpf 1984), Dutch (Housen 1994) and French (Noyau 1990, Trévise 1987)) and large-scale cross-sectional work (on English (Bardovi-Harlig 1992b, 1995, 1998, Housen 1998), French (Véronique 1987), German (von Stutterheim 1986) and Spanish (Lafford 1996)). By and large, the studies which have been conducted thus far have come out in support of the discourse hypothesis (but see Kumpf 1984 for a counter example), although there can be a fair amount of variation depending on the level of the learners. For instance, Véronique (1987) found that across levels there was a lot of variation in aspect marking with respect to foregrounding and backgrounding but that the emerging perfective morphology of his intermediate learners tended to appear in the foreground, while base forms dominated in the background.

---

[7]Confusingly, Bardovi-Harlig (1998, 2000) chooses to use terminology here that is specific to English, even though it is clear from her explanation of results from, for example, work on French that she in fact means that verbs in the foreground will carry *perfective* morphology. In her review of three recent books on the acquisition of second language aspect (Bardovi-Harlig 2000, Li & Shirai 2000, Salaberry 2000), Slabakova (2002) points out that the authors do not make a terminological distinction between past tense marking and perfective/imperfective aspect.

[8]Verbs in the background tend to be marked in the present tense or in a base form.

4.5 The Generative Studies

While the research reviewed thus far has sought to track the development of aspect among second language learners, work in the generative framework has used aspect to address broader theoretical questions about second language acquisition. As is typical of generative work in other areas of grammar, such as verb raising (White 1991, 1992, *inter alia*) and null subjects (Clahsen & Hong 1995, Hilles 1991, Liceras 1989, White 1985, 1986, *inter alia*), research has focused on whether second language learners maintain access to Universal Grammar (UG), and to what degree the first language plays a role[9]. They have also addressed, to a lesser extent, the issue of ultimate attainment.

In contrast to the majority of work in other approaches to aspect which exploits production data (Slabakova & Montrul 2002), generative studies of aspect are geared more towards comprehension. Specifically, researchers are interested in the extent to which learners are able to interpret the semantic properties of aspect in a native-like fashion. Studies of the perfective-imperfective distinction in Spanish (Montrul & Slabakova 2002, Slabakova & Montrul 2002) and telicity in English (Slabakova 1999) and Russian (Slabakova 2005) have shown that learners are capable of acquiring the semantic distinctions of the second language. Furthermore, in contrast to earlier work by Coppieters (1987), the competence of some near-native speakers is such that they are indistinguishable from natives (Montrul & Slabakova 2003).

Another interesting finding of the study by Montrul & Slabakova (2002), mentioned above, is that there were some intermediate learners who had mastered neither the aspectual morphology nor the semantic contrast while others were able to produce the morphology but were not sensitive to the semantic opposition. Furthermore, overall learners at intermediate and advanced levels who were competent in distinguishing aspectual meanings were also able to produce the morphology. This suggests that aspectual morphology emerges before learners understand what it means, analogous to Dietrich et al. (1995)'s observation that aspectual form precedes aspectual function (see section 4.2.3). However, the

---

[9]For the interested reader, three major hypotheses have been proposed on the issue of UG and first language transfer: the Partial Access Hypothesis (Bley-Vroman 1989, Clahsen & Muysken 1989, 1996, Schachter 1990, 1996), the Direct Access Hypothesis (Epstein et al. 1996) and the Full Transfer/Full Access Hypothesis (Schwartz & Sprouse 1996).

opposite was true in Slabakova (2005)'s study of anglophones learning how to mark telicity in Russian. She found that some low intermediate learners and most high intermediate and advanced learners were able to interpret telicity marking in Russian, but that they had difficulty in producing the requisite morphology. She attributed her results to the complexity of learning the lexical items necessary for expressing telicity in Russian.

A methodological weakness of several of the generative studies mentioned above is that they tend to use contrastive test items which facilitate the task. For instance, the sentence completion task used by Montrul & Slabakova (2002, 2003) and Slabakova & Montrul (2002) involved the presentation of a series of sentences which the learners were asked to rate on a scale from illogical (-2) to logical (+2). A series of verbs were chosen to represent accomplishments, states and achievements. For example, *vender* (to sell) was selected as an achievement. Each verb was then used in two separate sentences, one marked as perfective and the other as imperfective with minimal additional distinctions between sentences. For example (Montrul & Slabakova 2003, Appendix A) (with editing to indicate grammatical aspect):

(4.4)   *Los González **vendían**-IMP la casa pero nadie la compró.*
          The González's were selling their house but nobody bought it.
          *Mis padres **vendieron**-PERF el auto pero nadie lo compró.*
          My parents sold their car but nobody bought it.

The intention of the researchers was to create minimal pairs whereby only one type of grammatical aspect was target-like. However, even with distracters, the problem is that by contrasting grammatical aspect on the same verb, the semantic opposition under investigation is highlighted, which may give away the goal of the experiment or create a task effect.

4.6   Chapter Summary and Evaluation

The investigation of aspect in second language acquisition has been conducted from a number of perspectives, resulting in a varied contribution to the field. The descriptive studies, conducted under the auspices of the ESF project (Dietrich et al. 1995), identified three stages in the acquisition of aspect by second language

learners. Due to the absence of syntactic structure and morphology, the pre-basic variety, the first of the three stages, is characterized by the use of adverbial expressions (e.g. days of the week, specific years, time of day) and pragmatic means of expression (e.g. the principle of natural order, scaffolded discourse, implicit reference) to convey temporality. In the following stage, termed the basic variety, verbs are accompanied by their arguments and appear in a default form, along with a growing repertoire of adverbials and connectives. Observed by Dietrich et al. (1995) as the final phase observed in their learners, the further development stage is the point at which verbal morphology begins to emerge. This stage is characterized by slow, gradual and continuous progress; mastery of verb form before function; the emergence of irregular morphology before regular morphology and the marking of tense before aspect.

The aspect hypothesis studies have focused on the relationship between grammatical and lexical aspect. Specifically, they have found that perfective marking tends to begin on accomplishments and achievements, spreading to activities and states as learners develop, and that imperfective marking starts on states and activities, extending eventually to achievements and accomplishments. However, there has also been some suggestion that at the beginning stages of acquisition perfective past morphology is used as a default tense marker.

As the first step in a move towards determining the role of discourse in the acquisition of aspect, the discourse hypothesis states that learners use grammatical aspect to distinguish the foreground from the background in the structure of their narratives. Research conducted to date has come out in favour of the discourse hypothesis, subject to learner level.

Motivated by questions about access to UG and first language transfer, research thus far within the generative framework has found that learners are capable of acquiring the semantic properties of aspect in their second language. However, in some cases, task effects may have contributed to the positive results. Generative studies have also observed a pattern of morphology being acquired before meaning, although this may be subject to the inherent aspectual complexities of the second language.

The work reviewed above exhibits several shortcomings. Research on the aspect and discourse hypotheses is limited in that they isolate one factor in aspectual marking which leads to a partial picture of tense-aspect use (Bardovi-Harlig 2000). Furthermore, neither hypothesis explains the learning problem that aspect presents for the learner. Thus, while it may be true that perfective marking appears first on accomplishments and achievements, or that learners use grammatical aspect to distinguish foreground from background, without a theory of how aspect might be acquired, these types of learner behaviour cannot be explained. (For details on what we believe the acquisition of aspect entails, see Section 5.2.) There has also been an over-focus on production data, except in the generative studies and in a few of the aspect hypothesis experiments. Moreover, production data has predominantly taken the form of narratives, rather than exploiting other text types. More generally, the purpose of the vast majority of research to date has been to characterize the process of acquiring the tense-aspect system of a given target language. As yet, little attention has been given to studying whether external factors, such as instruction, can effect change in the route or rate of acquisition, despite specific calls for such research by Bardovi-Harlig (2000) and Salaberry & Shirai (2002).

In an effort to shift the focus of the current research agenda towards outside influences, our experiment will test feedback as an external factor which may have an effect on the acquisition of aspect. In doing so, we will also build on previous research by creating tasks that exploit more than one text type and which generate contextualized production and comprehension data.

CHAPTER 5

Methodology

5.1 Introduction

The foregoing chapters have revealed a need for interdisciplinary research on feedback in both ICALL (Intelligent Computer-Assisted Language Learning) and SLA (Second Language Acquisition) and have identified aspect as a suitable target structure on which feedback could be tested. More specifically, in chapter 2, we found from reviewing the ICALL literature that further quantitative testing of Intelligent Language Tutoring Systems (ILTSs) and the implementation of feedback informed by SLA research is necessary. In the same chapter, we also saw that the field of SLA would benefit from further empirical testing of feedback, particularly outside of the Interaction Hypothesis framework. The choice of aspect as a target structure was motivated in chapter 4. Our review of the literature on aspect in second language acquisition showed that the role of external factors, which include the effect of feedback, have not yet been considered in the process of acquiring aspect. It also illuminated two main shortcomings of previous research on the acquisition of aspect, namely that it has been over-focused on both narratives and production data.

The purpose of our experiment was to address these issues. Using a pre-test/post-test design, it sought to determine whether exposure to Input Processing (IP) or Explicit Inductive (EI) feedback delivered by an Intelligent Language Tutoring System (ILTS) would result in improved production and comprehension of past tense aspect in French among anglophone learners. Details of the feedback types

(Section 5.4), experimental design (Section 5.5), subjects (Section 5.7) and procedure (Section 5.8), as well as the statistical analysis and results (Section 5.9), will be presented in this chapter. We will also discuss the benefits of an ICALL approach for studying feedback in SLA (Section 5.3). We begin the chapter by defining the learning problem that French past tense aspect poses to Anglophone learners in order to establish what exactly needs to be acquired and how the feedback is intended to help.

5.2   The Learning Problem

As discussed in Chapter 3, aspectual meaning is composed of lexical aspect (i.e. a situation type, such as an achievement, which is itself composed of a verb and its arguments) and grammatical aspect. As a component in the composition of aspectual meaning, grammatical aspect imposes a particular viewpoint on a situation type (see Section 3.2.2). Representationally, we assume that there are separate categories of lexical and grammatical aspect. We consider the categories of lexical aspect (i.e. achievements, accomplishments, activities and states) to have the same features in French and English. For example, activities are considered [-static], [+durative] and [-telic] in both languages. Grammatical aspect, however, is distributed differently in French and English. Therefore, we assume that the categories of perfective and imperfective grammatical aspect exhibit different features. For instance, imperfective grammatical aspect in French must include features for habitual, progressive and non-progressive meaning while the equivalent category in English must comprise a feature for progressive meaning. Finally, we hypothesize that there are combinatorial rules for lexical and grammatical aspect which vary between languages. For example, a combinatorial rule for French would specify that states can combine with imperfective grammatical aspect. This particular rule would not be present in the grammar of an Anglophone, since English does not allow this combination of lexical and grammatical aspect.

The description of the French past tense aspectual system presented in Section 3.3.2 explained that the relationship between aspectual form and meaning is symmetrical. In other words, perfective verbal morphology expresses all situations as closed and imperfective verbal morphology conveys all situations as open. Thus,

to acquire French past tense aspect, second language learners must be able to interpret and produce any combination of situation type and perfective or imperfective grammatical aspect. We predict that aspectual meanings which are generated in the same way in French and English will not pose any problems to learners. For instance, to express progressive meaning in French and English, non-stative situation types and imperfective grammatical aspect are combined. For example, an accomplishment, such as [Ross climb a tree/Ross grimper dans un arbre], will combine with the morphology indicating imperfective grammatical aspect in French (*imparfait*) and English (*past continuous*) to denote an accomplishment that is in progress but has no endpoint (Ross was climbing a tree/Ross grimpait dans un arbre).

In contrast, the expression of non-progressive aspectual meaning, habituality and coercion are predicted to pose learning problems because aspectual meaning is not generated in the same way in both languages, owing to the differences in aspectual systems explained in Chapter 3. We hypothesize that the acquisition of these three structures by Anglophone learners will involve restructuring of the categories of grammatical aspect and the rules which govern the combination of lexical and grammatical aspect. We will explore each structure individually in order to identify the exact features that will need to be modified in the developing interlanguage grammar for acquisition to take place.

As discussed in Section 3.2.2 and reiterated above, French grammatical aspect has one category of imperfectivity which encompasses habitual, progressive and non-progressive meanings. English grammatical aspect, on the other hand, only conveys one imperfective meaning; namely, the progressive. Since progressive meaning is not compatible with states, it is ungrammatical to mark states with imperfective grammatical aspect in English (e.g. *Bill was knowing the answer*). It is possible in French, however, since, marked as imperfective, the state would convey non-progressive meaning[1]. We believe that to acquire non-progressive aspectual meaning in French, two conditions must be met by Anglophone learners. Firstly, they must extend the category of imperfective grammatical aspect in their interlanguage grammar to include non-progressive meaning. Secondly, they must alter their combinatorial rules such that states are licensed with imperfective grammatical aspect.

---

[1]Depending on the context, it could also convey habitual meaning.

We propose that similar restructuring of the interlanguage grammar is also required to acquire habitual meaning. English uses either *used to* or perfective morphology to express habituality. Recall from Chapter 3 that the status of perfective aspect in English is effectively neutral since the boundedness of a situation is determined by the category of lexical aspect to which the verb constellation belongs. Recourse is also made to the wider context both to exclude a bounded interpretation, where necessary, and to assign a habitual one. French uses imperfective morphology to convey habituality. Although descriptions of aspectual systems tend not to go beyond the sentential level, it must be the case that context disambiguates the three possible meanings denoted by imperfective grammatical aspect in French. Thus, again Anglophone learners must ammend their category of imperfective grammatical aspect to include habitual meaning. Moreover, they will have to modify their combinatorial rules to include imperfective marking indicating habitual meaning on all situation types.

As discussed in Section 3.4.3, states can be coerced into inchoative achievements in both French and English. States are exclusively marked with perfective morphology in English. Since states do not have endpoints and perfective aspect in English does not impose a bounded interpretation, coercion of states into inchoative achievements occurs via adverbials or context. French, on the other hand, simply applies perfective grammatical aspect. Therefore, to acquire coercion, Anglophone learners must shift the feature of perfective grammatical aspect from neutral to bounded.

5.2.1 *The Reliability of Context for Deriving Aspectual Meaning*

In principle, there is no reason that the learning problem posed by the target structures cannot be overcome. Learners should be able to come up with an interpretation from the context of the input and relate it to the aspectual form. Consider an example from the texts used for the experimental questions. This sentence is taken from an interview between Louis Malle and Philip French on the subject of the film *Lacombe, Lucien*[2]. At this point in the interview, Malle is describing Pierre Blaise, the actor who played Lucien in the film:

---

[2]To clarify, Philip French is the interviewer and Louis Malle is the interviewee and director of the film *Lacombe, Lucien*.

(5.1) Il était passionné de chasse; il parlait des oiseaux, les oiseaux à certaines saisons, comment les trouver, comment se cacher pour les tirer.

He was passionate about hunting; he talked/used to talk about birds, birds in particular seasons, how to find them, how to hide in order to shoot them.

Since this sentence describes Pierre's character, it is clear that the imperfective marking on the verb in the phrase *il était passionné de chasse* corresponds to a non-progressive imperfective meaning. Similarly, it follows from the fact that Pierre was passionate about hunting that he would have talked about birds repeatedly, rather than on a single occasion. Accordingly, the habitual imperfective meaning is marked on the verb *parler* with the *imparfait*. With sufficient exposure to input of this sort, the learner should be able to work out that the imperfective morphology on the verb corresponds to the imperfective meaning of the utterance. Notably, to achieve the desired meaning in this example, the possibility of using perfective aspect is excluded. With perfective marking, the verb phrases would become sequential, meaning that Pierre's passion for hunting would be felt at a particular moment and following that, his talking about birds would take place. Such a sentence would be ungrammatical and nonsensical. Unsurprisingly, the native speakers who participated in the experiment unanimously marked both verbs in this sentence with the *imparfait*.

Aspectual choice, such as that presented in Example 5.1, is not categorical in all contexts. Consider another example taken from the same interview. Here Philip French has just asked Louis Malle if it was true that Pierre Blaise was completely ignorant of what happened during the occupation of France during the second world war. Malle's reply in the transcript of the interview is:

(5.2) Oui. Il a fallu tout lui expliquer

Yes. It was necessary to explain everything to him (at a particular moment once and for all)

Although Malle's aspectual choice was the perfective, imperfective marking is also possible in this context. *Il fallait tout lui expliquer* means that the explanation happened gradually over a long period of time and did not include a specific endpoint. As noted in the English translation of the sentence in Example 5.2, implicit in the perfective meaning is that the explanation occurred at a specific

time in the past and had a clear endpoint. Analysis of the responses of the ten native speakers who completed the experimental questions confirmed that both aspectual markers were possible, as their choices were divided between the *passé composé* and the *imparfait*.

Example 5.2 raises an important point about the transparency of the perfective-imperfective aspectual contrast as it is expressed in the input available to learners. We hypothesize that the learnability of the perfective-imperfective aspectual contrast hinges on how transparently a meaning can be derived from the input. At one end of the spectrum are contexts in which only one of the aspectual meanings is licensed (see Example 5.1). We predict that the contrast as it is expressed in these contexts is learnable because the learner can recover the aspectual form-meaning relationship from the context. At the other end of the continuum are contexts in which both meanings are possible and along the continuum are contexts where native speakers tend to prefer one meaning over another. We expect that these cases will pose problems for learning the perfective-imperfective aspectual contrast. This is because if the context does not restrict aspectual choice to one marker or the other, the input will not allow the learners to retrieve consistent information about the relationship between aspectual form and meaning. We predict that restructuring of the categories of grammatical aspect and the combinatorial rules will therefore be impeded.

To our knowledge, there is no research which might illuminate how aspectual choice is distributed among native speakers. There is evidence from the native speaker data collected for the current experiment, however, that variation is quite common (see Section 5.9.1). For instance, analysis of the native speaker responses in the transformation exercises revealed unanimous aspectual marking in only 60% of cases. This suggests that the input contains many instances in which consistent expression of the perfective-imperfective form-meaning contrast is not available for learners. Intervention by feedback is one possible way of addressing instances where positive evidence will not be sufficient for the learner to induce the aspectual form-meaning relationship. Description of the feedback types and explanation of the ways in which they are intended to help the learners are presented in Section 5.4.

5.3   The Benefits of an ICALL Approach

The ability to handle free input has often been touted in the ICALL literature as one of the ultimate appeals of ILTSs. Consider, for instance, the *FreeText* project. One of its goals was to provide an ILTS which could analyze and correct complete sentences written by the learners in an effort not to be limited, as typical programmes are, by the structures that are hardcoded into the system architecture (L'haire 2004). Similarly, CASTLE (Krüger & Hamilton 1997, Murphy & McTear 1997) was conceived to be able to analyze and provide feedback on *almost* free input (Murphy & McTear 1997) in the context of communicative role-play scenarios. The main reason in favour of an ILTS which can analyze and correct any written learner utterance is that error correction would be exhaustive and individualized. While breadth of correction may be desirable from a pedagogical perspective, it is not appealing for experimental research on the effectiveness of feedback. Rather, the constraints imposed by computerizing a tutorial are what constitute the benefit of an ICALL approach to providing feedback. Firstly, the process of working through a computerized tutorial is very controlled. The content and presentation of instructions, the order of the tutorial questions and the content, amount and timing of the feedback is identical for all participants. Unlike with a human tutor, where there will always be slight variations in how instructions are explained or how feedback is communicated, the computerized tutorial ensures that every learner has the same experience.

Computerization also constrains the task. In the case of our ILTS, there was a transformation task to test production of aspect and a grammaticality judgment task to evaluate comprehension of aspect (see Section 5.6.1 for details of both tasks). While comprehension tests for specific areas of grammar are inherently constrained, production tasks can range widely in how restrictive they are. Gap-fills, for instance, are highly controlled, while compositions are not. In an ILTS, a controlled production task is necessary in order to guarantee accurate error identification and feedback (see Section 5.6.1 for further discussion of this point). Presenting constrained tasks minimizes the influence of grammatical considerations or problems beyond those of the target structure. This is advantageous for testing the effectiveness of feedback because it will strengthen the claim that any learning gains are attributable to the feedback.

Finally, the ILTS may help in addressing the "blame-assignment" problem (we refer the reader back to Section 1.1.3 for clarification). To ensure that the tutoring system will provide accurate feedback, learner errors must be predictable. To make the errors predictable, the task has to address a single target structure in a controlled fashion. Through the computerization of the task and the feedback, the work of identifying which rule in the interlanguage grammar needs to be changed or abandoned has already been significantly reduced for the learner. This increases the possibility that he or she will succeed in finding and addressing the faulty rule.

5.4   Establishing Feedback Types

5.4.1   Explicit Inductive (EI) Feedback

As discussed in Sections 2.3.4 and 2.4.1, explicit inductive (EI) feedback was identified as one of two types that would be worth testing. The idea was that it would be explanatory, following the success of work on deductive feedback by Carroll (2001), Carroll & Swain (1993) and Kubota (1997). However, rather than explaining how the learner response violated the rules governing the target structure, EI feedback would explain what each non-target-like response means. In this way, the learners would be encouraged to induce what aspectual meanings are not associated with perfective and imperfective aspect and, by extension, which ones are.

EI feedback conformed to the following structure: *"What you wrote means X*[3]. *This does not make sense according to the context*[4]*."* For a contextualized example, see Figure 5.5.

5.4.2   Input Processing (IP) Feedback

Input Processing (IP) feedback was inspired by the work of Bill VanPatten on his approach to grammar teaching called Processing Instruction (PI) discussed in

---

[3]The meanings contained in all the feedback messages were informed by the native speakers of French who participated in the experiment.

[4]For cases in which both the *passé composé* and the *imparfait* were possible but expressed slightly different meanings, the feedback message read, *"This is possible but not the most natural meaning according to the context."*, rather than *"This does not make sense according to the context"*.

Section 2.5. The crucial dimension of PI, as identified in studies by Benati (2004), VanPatten & Oikkenon (1996) and Wong (2004), is its structured input activities. In these activities, the learners are presented with either a picture or a sentence in the target language and are asked to match it to one of two possible meanings. One is the target meaning, while the other is the meaning resulting from a faulty input processing strategy predicted by VanPatten's model of input processing. The learners are then told whether they made the correct match (we refer the reader back to Example 2.21 for a sample structured input activity).

Taking the learner error as its starting point, IP feedback involved two parts. In keeping with the design of structured input activities, the learners were first presented with two interpretations: one of the response they had given and the other of the correct response. They were then asked to choose the interpretation that was most appropriate according to the context. The second part of the IP feedback informed the learners whether the interpretation they selected was appropriate or not. If it was, this was affirmed and they were told that their response conveyed the other interpretation. If it was not, they were told that the other interpretation was more natural. For a specific example, see Figures 5.7 and 5.8.

5.4.3   How the Feedback is Intended to Help

In Section 5.2, we hypothesized that in order to acquire aspect, learners must be able to restructure both the features of the categories of perfective and imperfective grammatical aspect and the content of the rules which govern how lexical and grammatical aspect are combined. We also proposed that the success of this restructuring is contingent upon how transparent the aspectual meaning is in the input. In particular, if a single aspectual meaning is licensed in the input, inducing the aspectual form-meaning relationship should be straightforward. However, if more than one aspectual meaning is possible according to the context, the form-meaning relationship becomes difficult to determine on the basis of positive evidence alone. We believe that in these cases feedback could be used to clearly define the aspectual meaning and relate it to the relevant aspectual form, allowing restructuring to proceed. Furthermore, in instances where aspectual choice is categorical, feedback may expediate the restructuring of the Anglophone learner's interlanguage aspectual system.

Both EI and IP feedback focus on explaining the aspectual meanings associated with perfective and imperfective verbal morphology. In our view, explicitly stating the meaning conveyed by a perfective or imperfective form will clarify what features need to be present in their category of grammatical aspect. It will also inform learners about the rules surrounding what combinations of grammatical and lexical aspect are licensed, or not, in the target language. For instance, we predict that learners will erroneously use perfective marking on states to express non-progressive meaning. In response to this error, EI feedback tells the learner what this combination of grammatical and lexical aspect means in French. Since there is only one other aspectual marker in the past tense, the learner should be able to conclude that one of the features of imperfective grammatical aspect is to express non-progressive meaning. It should also be clear that imperfective grammatical aspect can combine with states.

Addressing the same error, IP feedback presents two meanings: the meaning resulting from the combination of perfective aspect and state (i.e. the learner's response) and the meaning resulting from the combination of imperfective aspect and state (i.e. the target-like response according to the context of the question). The learners are then asked to think about what is happening at this point in the text/interview and to select the interpretation which is most natural in the context. They are then given feedback on their choice.

If, in this example, the learners select the perfective interpretation, the ILTS will tell them that the other interpretation is more natural according to the context. Selection of the non-natural interpretation presents a different problem from that which the feedback is intended to address. Specifically, we believe that it is the expression of aspect in French, rather than the aspectual meanings themselves, which pose problems to learners. In other words, the same aspectual meanings are available in French and English; they are simply expressed differently. Therefore, if the complications of form are factored out of the equation, the learners should have no difficulty in selecting a target-like aspectual meaning. In the unlikely event that they do select a non-natural aspectual meaning, it most likely means that they have misunderstood the text.

On the other hand, selection of the non-progressive interpretation, in this example, is evidence of the predicted mismatch between aspectual meaning and as-

pectual form. That is, it shows that the learners understand that non-progressive aspectual meaning is appropriate in the context, but, on the basis of their response, have used perfective morphology to express it. In this case, the learners are informed by the ILTS that they have chosen the appropriate interpretation, but that their response conveys the other interpretation (i.e. the perfective one). Learners should be able to conclude from this feedback that perfective marking of states does not result in non-progressive aspectual meaning in French. They should also be able to infer that in French imperfective grammatical aspect expresses non-progressive meaning on states. On the basis of this information, restructuring of the learner's category of imperfective grammatical aspect should occur to include non-progressive meaning. This information should further allow the modification of their combinatorial rules such that imperfective grammatical aspect marking is licensed on states.

5.5   Experimental Hypotheses and Design

The experiment sought to test whether EI and IP feedback are effective in helping anglophone learners to acquire French past-tense aspect. The hypotheses were as follows:

- **Hypothesis 1**: Competence in the production and comprehension of past-tense aspect among learners who receive EI feedback will increase, as measured by differences in pre and post test scores, compared to those who receive no feedback.
- **Hypothesis 2**: Competence in the production and comprehension of past-tense aspect among learners who receive IP feedback will increase, as measured by differences in pre and post test scores, compared to those who receive no feedback.

The hypotheses were tested using a pre-test/post-test design involving two experimental groups and one control group as follows:

- **EI Group**: The learners wrote a pre-test, were exposed to EI feedback in response to non-native-like answers and congratulatory feedback in response

to native-like answers during the experimental treatment, wrote one post-test immediately following the experimental treatment and a second post-test three weeks later.

- **IP Group**: The learners wrote a pre-test, were exposed to IP feedback in response to non-native-like answers and congratulatory feedback in response to native-like answers during the experimental treatment, wrote one post-test immediately following the experimental treatment and a second post-test three weeks later.

- **Control Group**: The learners wrote a pre-test, were exposed to no feedback, neither corrective nor congratulatory, during the experimental treatment, wrote one post-test immediately following the experimental treatment and a second post-test three weeks later.

5.6   Materials

5.6.1   Exercise Types

The vast majority of the literature to date on the acquisition of aspect by second language learners has focused on production, rather than comprehension. The ESF studies (Dietrich et al. 1995), most of the work on the aspect hypothesis (Andersen 1986, 1991, Bardovi-Harlig 1992a, 1998, Bardovi-Harlig & Bergström 1996, Bardovi-Harlig & Reynolds 1995, Bayley 1994, Bergström 1995, 1997, Cadierno 2000, Collins 1997, Comajoan 1998, Flashner 1989, Giacalone Ramat 1995, 1997, Giacalone Ramat & Banfi 1990, Hasbún 1995, Housen 1993, 1994, Kaplan 1987, Kihlstedt 2002, Kumpf 1984, Leiria 1994, Leiria & Mendes 1995, Ramsay 1990, Robison 1990, Rohde 1996, Salaberry 1999, Wiberg 1996) and the discourse hypothesis studies (Bardovi-Harlig 1992b, 1995, Housen 1998, Lafford 1996, Noyau 1990, von Stutterheim 1986, Trévise 1987, Véronique 1987) have looked exclusively at production. Only generative work has investigated the comprehension of aspect (Montrul & Slabakova 2002, 2003, Slabakova 1999, 2005, Slabakova & Montrul 2000, 2002) and very few studies have looked at both comprehension and production together (Salaberry 1998, Shirai & Kurono 1998). Given the widely-held observation that comprehension tends to precede production in language acquisition, it would be interesting to know whether feedback affects both simultaneously or one before the other. To shed some light on this question and

to redress the imbalance in the literature towards research that focuses only on the production of aspect, one exercise type to test production and another to test comprehension were deemed worthwhile for the experiment.

The Production Task

Of the studies on the acquisition of aspect which have focused on written competence, production data have been elicited using two main techniques: cloze passages (Bardovi-Harlig 1992a, Bardovi-Harlig & Reynolds 1995, Bergström 1995, 1997, Collins 1997, 1999) and written narratives retelling film storylines[5] (Bardovi-Harlig 1998, Bardovi-Harlig & Bergström 1996, Bergström 1995, 1997, Hasbún 1995, Salaberry 1998). Although cloze passages and retell tasks are specific techniques, their advantages and disadvantages embody the trade-off in the experimental design between implementing tasks controlled by the learner versus those controlled by the experimenter. Like all opened-ended exercise types, retell tasks are advantageous because they allow for learner-controlled language production which, by virtue of being based on the same content, can, in theory, be compared across learners. However, in practice, retell tasks result in a large amount of variation in the number and type of tokens produced by individual learners, rendering any systematic comparison problematic (Bardovi-Harlig 2000). Cloze passages, on the other hand, allow for straightforward comparisons between subjects, strict experimenter control over what aspects of language are elicited and the reduction of ambiguous learner production, but are rather contrived and mechanical as well as limited in the scope of the learner's language production competence that they are able to capture.

For the purposes of this experiment, the choice of production task was restricted by the target structures under investigation and by computational considerations. Three specific manifestations of aspect were identified for testing based on our discussion of the differences between the French and English aspectual systems in chapter 3 (see, in particular, Section 3.5): the *imparfait* for expressing habituality, the *imparfait* for conveying non-progressive continuous meaning and aspectual coercion. In order to ensure that the learners produced all three target

---

[5]This technique has also been widely used to collect spoken data. (See Bardovi-Harlig (1998), Comajoan (1998), Giacalone Ramat (1997), Lafford (1996), Liskin-Gasparro (1997) and Salaberry (1999) for examples.)

structures and that there was enough opportunity for them to do so, the production task had to afford sufficient control over the language produced by the learners. This was particularly important for testing coercion, as occurrences of it are typically infrequent and highly marked.

Computationally, the priority was for accuracy in error identification and subsequent feedback delivery by the tutoring system. This was paramount for instilling confidence in the user with regard to the capability of the system and could only be guaranteed by using a production task which allowed a finite number of predictable learner responses. This point is taken up in detail specific to the system developed for the experiment in Section 5.8.2 under *Instructions*.

Taking into account the advantages and disadvantages of production tasks used in previous research and the restrictions posed by the target structures and the computational considerations of the current study, a transformation exercise was selected as the means for eliciting the production data. For this task, the learner is presented with a text in the present tense and is asked to "transform" it (i.e. rewrite it) into the past tense. As such, a transformation exercise afforded sufficient control over the number and distribution of the target structures of aspect and facilitated successful analysis of the learner input by the tutoring system. Admittedly, the transformation exercise suffers from the same disadvantages as the cloze task in that it is contrived and mechanical, as compared to a learner-controlled task (e.g. a composition), and is therefore less informative as regards the learner's overall competence in written production. However, these disadvantages are perhaps less pronounced in the transformation task in that it encourages the learner to view the text as a whole and to consider each sentence as part of a larger context.

The Comprehension Task

Flavours of comprehension task that have been used in previous studies on the acquisition of second language aspect include a cloze passage (Salaberry 1998), a multiple choice acceptability judgment test (Shirai & Kurono 1998), a sentence-conjunction judgment task and a truth-value judgment task (Montrul & Slabakova 2002, 2003, Slabakova & Montrul 2002).

The cloze passage, as used to test comprehension by Salaberry (1998), was a revised interview taken from Paris Match magazine in which the learners chose between the *imparfait* and *passé composé* forms of the verb provided for each question. While worthwhile in that the results could be compared to the production data that was collected as part of the same study, Salaberry's cloze passage is problematic in two respects: firstly, it does not include any distractors. Consequently, the learners could have conceivably deduced that the study was testing the aspectual distinction between the *imparfait* and the *passé composé* and they had a 50/50 chance of making a native-like judgment by simply guessing. Either possibility could have biased the results. Furthermore, a task that encourages the learner to make categorical choices is perhaps not always the most appropriate and, in some cases, may prove to be inaccurate for testing the acquisition of aspect. While it appears that in Salaberry's particular interview the aspectual choices might have been fairly clear, this would be subject to confirmation by the individual native-speaker data, which is not reported in the study. Certainly in the current study there were many instances in which the native speakers did not make a unanimous aspectual choice.

For the multiple choice acceptability judgment test, devised by Shirai & Kurono (1998), the learners were given a series of short dialogues, each with a missing verb phrase. Two to four possibilities for the missing verb phrase were provided and the learners were invited to judge each one as *correct*, *incorrect* or *not sure*. Notably, they could judge more than one possibility as correct. This approach is more flexible than the cloze passage in that it offers the learner more than two choices; however, like the cloze passage, the choices are still categorical. Our data indicate that in contexts where two aspectual markers are possible, native speakers tend to prefer one over the other. Accordingly, comprehension tasks for testing the acquisition of aspect should model native speaker behaviour in this respect.

Reused in several studies by the same authors (Montrul & Slabakova 2002, 2003, Slabakova & Montrul 2002), the sentence-conjunction judgment task entailed presenting the learners with two clauses coordinated with either the conjunction *and* or the conjunction *or* (we refer the reader back to Example 4.4 for two sample questions). The learners were then asked to judge whether each sentence made sense on a scale of -2 (illogical) to +2 (logical). Their responses would show

	coercion	habitual	non-progressive	distractors
Transformation Interview	3	1	4	5
Transformation Narrative	0	6	1	12
Total for Transformation	3	7	5	17
GJ Interview	8	4	4	13
GJ Narrative	2	1	0	7
Total for GJ	10	5	4	20
Grand Total	13	12	9	37

Table 5.1: Distribution of target structures by exercise type

whether or not they understood the bounded interpretation of the preterite versus the unbounded interpretation of the imperfect. Although having the learners express aspectual preferences by ranking was appealing, using conjoined sentences was not conducive to testing the manifestations of aspect that we were interested in, as they could not be presented in opposition to each other. The same was true for the truth-value judgment task. For it, the learners read a series of unrelated short stories, each of which was followed by a statement pertaining to the story which expressed either a bounded or an unbounded interpretation. The learners' task was to judge whether the statement was true or false.

To keep the learners from working out the purpose of the experiment and to reduce the likelihood of correct answers resulting from guesses, it was decided that the comprehension task should be a grammaticality judgment exercise in which the learners were presented with four possible past tenses, namely the *imparfait*, *passé composé*, *plus-que-parfait* and *conditionnel passé*. To reflect the fact that aspectual choices are not necessarily clear-cut and to assess whether or not the learners' performance approximated that of the natives in this regard, the learners were asked to rank how natural each of the four tenses were according to the context.

The distribution of target structures and distractors by exercise type are given in Table 5.1. While ideally there should have been an equal number of occurrences of each structure, it was not possible to achieve this without sacrificing the coherence of the text. There were a total of 32 Transformation questions and 40 Grammaticality Judgment (GJ) questions. Details of exactly what was required to complete each task are given in Section 5.8.2 under *Tasks*.

5.6.2 Programme Design

The tutoring system used to carry out the experimental treatment was built entirely in Python using cgi-scripts. It did not rely on natural language processing techniques, such as parsing or chunking, as the necessary linguistic analysis could be achieved using a combination of Python's built-in functions and the Association des Bibliophiles Universels's morphological dictionary of over 300,000 common French words[6]. The generation of the feedback was not automatic, as the content of every feedback message was unique to each question. As such, messages for all possible tense errors were stored in a series of text files, one for each set of questions.

Learner Input Analysis

Learner input came in the form of either numbers, for the comprehension task (see Section 5.6.1 above) or strings, for the production task (see also Section 5.6.1 above). The analysis of the numerical input was straightforward. It was simply a matter of matching the sentence which the learner marked as "1" with its tense, looking up the pre-stored correct answer and checking it against the student input and, as required, generating a feedback message appropriate to the feedback group the learner belonged to.

The analysis of the strings, on the other hand, involved a couple of stages of processing. Firstly, the verb had to be identified and extracted from the sentence that the student entered[7]. This was achieved using *SequenceMatcher*, an operator built into the *difflib* module in Python, which matches the characters of two strings. A pre-stored string of the correct answer with the verb delimited in square brackets was matched against the student response. Since conjugation and tense errors were to be expected in the student response, the five characters on either side of the square brackets in the correct answer were matched against the student's string[8]. The collection of characters in between were identified as the verb and extracted for analysis.

---

[6]This dictionary, as well as several others, are freely available for download from the Internet at http://abu.cnam.fr/DICO.

[7]I gratefully acknowledge Bill Winder for writing this module of the tutoring system for me.

[8]The obvious problem with this approach is that it relies on the five characters on either side of the verb in the student response being correct. Ultimately, this posed no problem to the tutoring system, for reasons discussed in Section 5.8.2 under *Instructions*

After checking that the verb did not match the pre-stored correct answer, stage two involved identifying the tense of the verb entered by the student. First of all, the extracted verb was classified as either a compound tense (i.e. composed of an auxiliary verb and a past participle) or a simple tense (i.e. a single verb made up of a stem and an ending) using a regular expression. The regular expression identified the extracted verb as a compound tense by checking to see if it matched the pattern of any character, followed by whitespace, followed by any character. If it did not match this pattern, it was assumed to be a simple tense. Further analysis was then required to identify the exact tense of the extracted verb. This was done by searching for the extracted verb in a morphological dictionary of French and returning the tense feature or, in the event of conjugation errors, matching the extracted verb against a verb stem.

The tutoring system could also handle various types of anomalous user behaviour. These are discussed in Section 5.8.2 under *Instructions*.

Special Features

A glossary and a conjugator were included in the design of the system so that the learners' aspectual choices would not be hindered by lack of familiarity with particular verb forms or lexical items. The glossary comprised an alphabetized list of words that we predicted may be unknown to the learners. The conjugator provided the forms of the imperfect and the past participle of the verb for the current question as well as the present, imperfect and present conditional forms of *être* and *avoir*. Both were displayed in pop-up windows which the learner could access as required by clicking on the relevant link.

5.6.3   Text Types

To be able to generalize the results across more than one text type and to exclude the possibility that aspectual choice might be influenced by text type, both interviews and narratives were incorporated into the experimental treatment.

As the experiment was originally targeted at French 1 A and B students from the French Section of the Division of European Languages and Cultures at the University of Edinburgh, the interviews and narratives chosen for inclusion in the experimental treatment were selected so as to be relevant to the course. As such,

	coercion	habitual	non-progressive	distractors
Transformation Interview	3	1	4	5
GJ Interview	8	4	4	13
Total for Interviews	11	5	8	18
Transformation Narrative	0	6	1	12
GJ Narrative	2	1	0	7
Total for Narratives	2	7	1	19
Grand Total	13	12	9	37

Table 5.2: Distribution of target structures by text type

the interviews were with the directors of the films which the students watched on the course. One was with Claire Denis, director of *Chocolat* taken from the Autumn 2001 edition of the French magazine *Modam* (Mon oncle d'amérique). The other was with Louis Malle, director of *Lacombe, Lucien* who published a series of interviews he gave about his films (Malle 1993). The narratives were excerpts from the books and short stories used on the course, specifically *L'Etranger* (Camus 1996) and *La Cérémonie* from the collection of short stories entitled *Jazz et vin de palme* (Dongala 1996). As the subjects who ultimately formed the sample were *not* French 1 A or B students, a brief introduction to each text was included so as to ensure that the excerpts would make sense to people who had not seen the films or read the books.

The distribution of target structures and distractors by text type are illustrated in Table 5.2.

5.6.4   Pre and Post Test Design

Isomorphic pre and post test tasks were devised to directly test the comprehension and production of aspect among the subjects in a way which was different from that of the experimental treatment (i.e. not a transformation or grammaticality judgment task) to allow the results to be generalized beyond the requirements of the experimental tasks[9]. At the same time, they were designed to provide contexts which evoked the target uses of aspect specifically and which

---

[9]It would have also been desirable to have conducted a more open-ended test of the subjects' aspectual competence (e.g. a composition) in order to determine whether or not the feedback had had a further generalizable effect. However, due to the length of the experiment and the specificity of the target structures, it was not possible to include any such additional assessment.

asked questions that could only be answered in a target-like fashion by drawing on sound aspectual knowledge. With this in mind, contexts that were one to five sentences in length were created based on excerpts from *Lacombe, Lucien* and *Chocolat* as well as several editorials from the *Femmes D'Ailleurs* column of *Le Courrier International* online within which the target structures could be incorporated.

The task to assess comprehension consisted of completing a sentence within the context provided by choosing between two possible answers. In every case, the two choices were contrived such that only one of them was natural in light of the aspect of the sentence. For instance, in Question 6 of comprehension Test B reproduced below (see Appendix A), because the verb *envoyait* is in the imperfect tense, a habitual interpretation is evoked and this matches most naturally with (a):

> *La Gestapo envoyait un petit cercueil* ⸺⸺⸺⸺⸺⸺⸺⸺⸺
> The Gestapo used to send a miniature coffin ⸺⸺⸺⸺⸺⸺
>
> *(a) aux familles de leurs prochaines victimes.*
> (a) to the families of their next victims.
> *(b) à la mère de Lucien.*
> (b) to Lucien's mother.

To evaluate production, the contexts were accompanied by an instruction to describe or explain something specific or make a particular statement relevant to the context. In all questions, part of the intended answer was already given in order to limit the scope of possible responses, thereby facilitating the elicitation of the target structure. In the example given below (Question 8 of production Test A, see Appendix A), the question prompts the coercion of a stative predicate *elle avoir peur des hyènes* (she to be afraid of the hyenas) into an event by the use of the *passé composé* (i.e. *elle a eu peur* (she was afraid)):

> *Un soir France entend des hyènes près de la maison. Elle court à la chambre*
> *de sa mère. Armé d'une mitrailleuse, Protée prend garde aux hyènes à la*
> *fenêtre de la chambre. France se calme.*

One evening, France hears hyenas near the house. She runs to her mum's bedroom. Armed with a machine gun, Protée stands guard against the hyenas at the bedroom window. France calms down.

Describe France's emotions upon hearing the hyenas:

Soudainement, elle _____
des hyènes.
Suddenly, she _____
of the hyenas.

As the above example also shows, the contexts were written in the present tense and the instructions were given in English, consisting of an imperative followed by noun phrases and prepositional phrases in most cases. These measures were taken so as to avoid exposing the subjects to aspectual input, thereby eliminating the risk of priming effects.

In each of tests A, B and C, there were 18 target items (9 comprehension and 9 production, each of which included three questions per target structure) and 4 distractors (2 comprehension and 2 production). The naturalness of all the expected responses with regard to aspectual choice was subject to confirmation by 10 native speakers. These data will be discussed in detail in Section 5.9.1.

5.7   Subjects

Recruiting subjects proved to be a difficult and time-consuming undertaking. As mentioned in Section 5.6.3, the texts for the experimental tasks were chosen with French 1 A and B students in mind as subjects. The idea was to have them participate in the experiment immediately preceding the spring holiday so that they could then complete the delayed post test without having had any exposure to French in the interim. Students who were exempt from the final exam were solicited directly by email to avoid running subjects who would need to study during the holiday. The experiment was also advertised at the University of Glasgow, St. Andrew's University, Heriot-Watt University and the University of Leeds. We hoped that by targeting sizable first-year classes with the help

of the course coordinators, we would be able to attract a large number of subjects[10]. Unfortunately, the response was abysmal, necessitating a reduction in experimental questions by 70% from 244 questions to 72.

A second attempt at recruiting subjects targeted learners who were at the French 1 A and B level (i.e. they had completed the high school prerequisites), but had chosen not to take French at university. They were all anglophones who were recruited at the University of Edinburgh from first-year courses in the subject areas of Archaeology, Architecture, Astronomy, Canadian Studies, Chemistry, Criminology, Divinity, Economics, Engineering, English Literature, Informatics, Linguistics and English Language, Mathematics, Philosophy, Physics, Politics, Psychology, Music, Nursing, Social Anthropology, Social Policy and Sociology as well as via an advertisement on the university's Careers Service website. There were a total of 29 participants[11], 11 male and 18 female, ranging in age from 17 to 20[12]. All participants were paid £10 for completing the experiment and were offered a selection of chocolates at the end of the immediate post test.

5.7.1 Subject Profile Questionnaire

To confirm the homogeneity of the sample, information which could have influenced subject performance was collected via an online subject profile questionnaire completed by every subject before beginning the experimental treatment. This included information on the following points:

- age of first exposure to French
- length of period of study
- details of formal training
- any extra-curricular activity and the frequency of it
- the number and duration of any holidays in a French-speaking country
- *current* knowledge of any second languages other than French and the duration of formal training

---

[10]The number of students taking French 1 A and B at Edinburgh alone who fit our criteria totaled 95.

[11]There were an additional three subjects who completed the experiment but were ultimately excluded for reasons discussed further on in this section.

[12]Parental consent was obtained for the one participant who was under the age of 18.

- *previous* knowledge of any second languages and the duration of formal training as well as current frequency of usage
- bilingualism and frequency of usage
- the duration of any period of residence in a French-speaking country and age at the time of arrival

Questions regarding the facility with which the subjects used computers were not asked on the assumption that the amount of computational skill required was too minimal to cause any distraction. Furthermore, during the experiment, all of the subjects were clearly very comfortable operating in a computational medium.

The age of first exposure ranged from 3 to 12, averaging 8.8 years. Almost all subjects had received their training in the British school system over the course of an average of 7.3 years and had attained A levels, Highers or Advanced Highers in French. None of the subjects was studying French at the time of the experiment and had last been exposed to the language regularly an average of 10 months previously. All of the participants had either current or previous knowledge of other second languages but they were languages that had been learned at school or beginner courses at university and none claimed to be bilingual in any of them. Similarly, the amount of time spent in a French-speaking country, either on holiday or in residence, was considered minimal. In light of this information, the sample was considered homogeneous with respect to these external variables[13]. Further details of the responses to the questionnaire can be found in Appendix B.

Three subjects from whom data had been collected were excluded from the analysis because one of their responses on the questionnaire revealed a source of variation which would have compromised the homogeneity of the sample. Specifically, one of them mentioned having a francophone flatmate for two years, another had a francophone parent and the third had lived in France until the age of twelve.

---

[13] Admittedly, the age of first exposure, the number of years of formal education in French and the amount of time without regular exposure to French before the experiment showed quite a range of variation across learners. To confirm that none of these three variables influenced the results of the experiment, a post-hoc regression analysis was carried out which compared each variable to the gain scores across all learners. The results showed that none of these variables correlated with gain scores (age of first exposure: $F_{(27,1)}=.532, p > .05$; number of years of formal education: $F_{(27,1)}=.256, p > .05$; time without exposure: $F_{(27,1)}=.108, p > .05$). Therefore, we can be confident that they did not influence the results of the experiment.

5.8 Procedure

5.8.1 Determining Level: The Placement and Pre Tests

Knowing that course level (i.e. A level, Higher or Advanced Higher) alone would allow too great a difference in tense-aspect use across learners (Bardovi-Harlig 1992c), the first step in the experimental procedure was to assess the subjects' level of proficiency in French. A placement test[14] and a pre-test (see Appendices A and C for exemplars) served this purpose. The placement test was administered first and evaluated the subjects' general grammatical knowledge of French. The test items included twenty gap-fill questions on present and past tense conjugation, question formation using inversion, prepositions, adjective and determiner agreement, pronouns and usage of the subjunctive mood. The number of questions and value of the points for each area of grammar were weighted in the original test in order to place more importance on some areas of grammar (e.g. verb conjugation) than others (e.g. pronouns) (see Appendix C for details). This marking scheme was basically maintained, but was subject to a finer gradation in that half the value of the question, rather than the whole value, was deducted for errors of gender, person and number agreement[15].

After completing the placement test, subjects were given test A, B or C as a pre-test which was completed immediately before the experimental treatment[16]. In addition to providing a benchmark from which improvement would be measured, the pre-test scores served as further confirmation of a uniform level of

---

[14]The placement test belonging to the Department of French, Hispanic and Italian Studies at the University of British Columbia was used with permission. Undergraduate students in that department who are not products of the local school system take the test so that they can be admitted to the appropriate course for their level of proficiency. The French Section of the Division of European Languages and Cultures at the University of Edinburgh have their own language test which is used to identify students taking French 1A who have trouble with basic grammar. Since at the outset students from French 1A were to be a significant part of the target audience for the experiment, it was not possible to use this language test.

[15]This was *not* the case was for the questions on adjective agreement, as the adjectives were supplied.

[16]The order in which tests A, B and C were administered was not randomized, as, due to the small number of experimental participants, this would not have guaranteed an even distribution of the tests. Rather, one of the six possible combinations of tests A, B and C (i.e. ABC, BCA, CBA, ACB, BAC or CAB) was assigned to each subject who then received the pre, immediate post and delayed post tests according to the combination order (i.e. If combination ABC was assigned to subject 1, he or she would write test A as a pre test, B as an immediate post test and C as a delayed post test).

proficiency among the subjects, specifically on their knowledge of past tense aspect.

5.8.2 The Experimental Treatment

The experimental treatment was carried out on personal computers on the Linguistics and English Language network in the experiment room of the Adam Ferguson Building on the George Square campus at the University of Edinburgh. Each participant was allocated to a booth which housed a computer, desk and chair. The doors to the booths were closed over during the experiment to avoid distraction but not shut completely so as to allow for airflow. Entirely web-based, the experimental treatment included a total of 72 questions shared across 4 texts, each text comprising a set of questions. To reduce fatigue effects, the subjects were given a break after the second set of questions.

The first task was to complete the subject profile questionnaire (see Section 5.7.1 above and Appendix B for details). At the end of the questionnaire, the subjects entered either their matriculation number or another form of unique identification of their choice as a login which was used by the system to record all of their responses.

Instructions

The next page, reproduced below in Figure 5.1, explained what they could expect while participating in the experiment. In addition to information about the number of questions and the exercise types, the subjects were also told on this page that if they made a mistake, they would receive a feedback message which would address mainly their choice of tense, rather than the correctness of their conjugations. This point was made explicit to mitigate the tension between reducing cognitive load and correcting erroneous input. Since aspect poses such difficulties to learners, it was considered desirable for the subjects to focus their attention on meaning rather than on form in order to reduce their cognitive load and maximize the possibility that the feedback would be processed. However, from a pedagogical perspective, it would have been undesirable for form errors
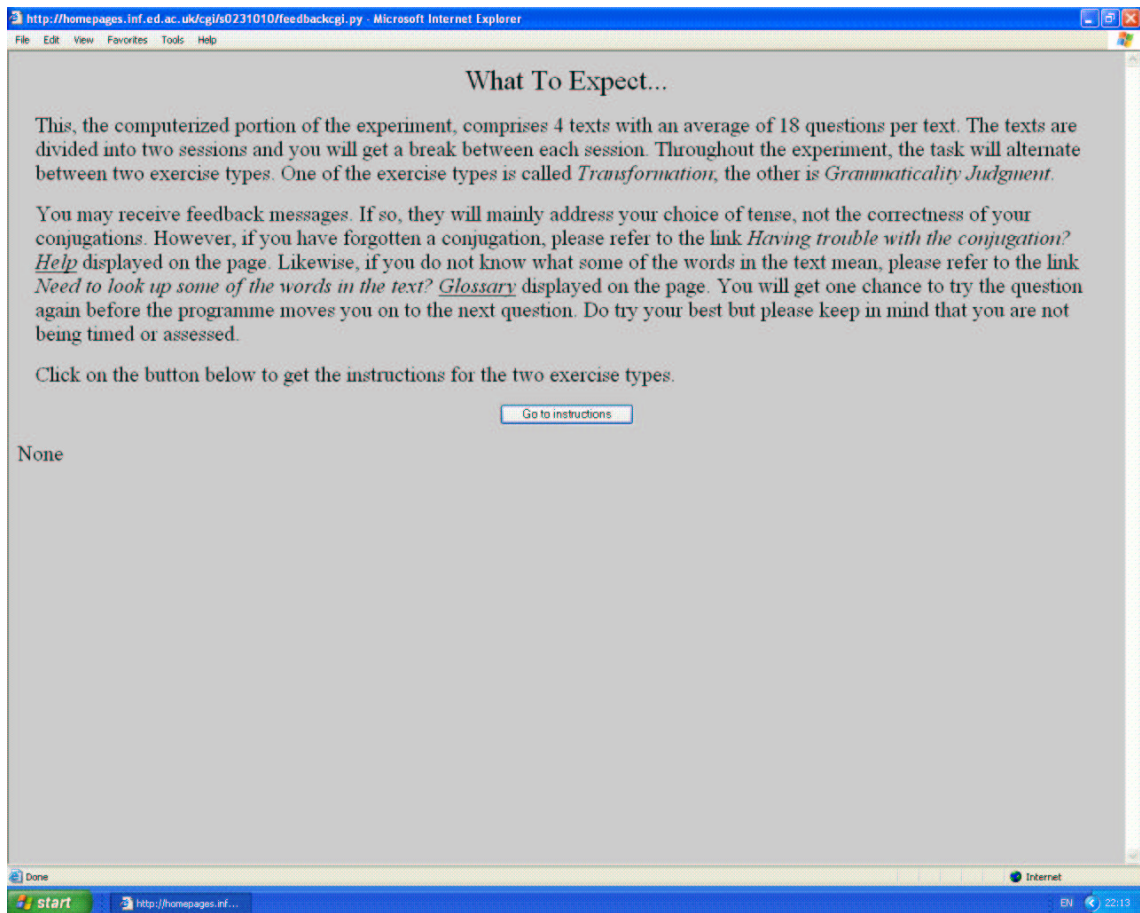
Figure 5.1: Screenshot of the page explaining what to expect during the experiment

to have been ignored completely, as this may have encouraged their perpetuation. The content of the feedback which addressed form errors is discussed in Section 5.8.2 under *Tasks*.

Also worth mentioning is the point about students being allotted two attempts at each question before being moved on to the next question. The idea behind this design decision was to control for a possible effect from repetition of the feedback message and to prevent a situation in which the learners could not figure out the right answer and became trapped in a frustrating loop of being told that they were wrong without being able to move on to the next question. The students were alerted to the fact that they would be limited to two attempts at each question so that they would know from the start that if they wanted to get the answer right, they would have to pay attention and consider their answer carefully.
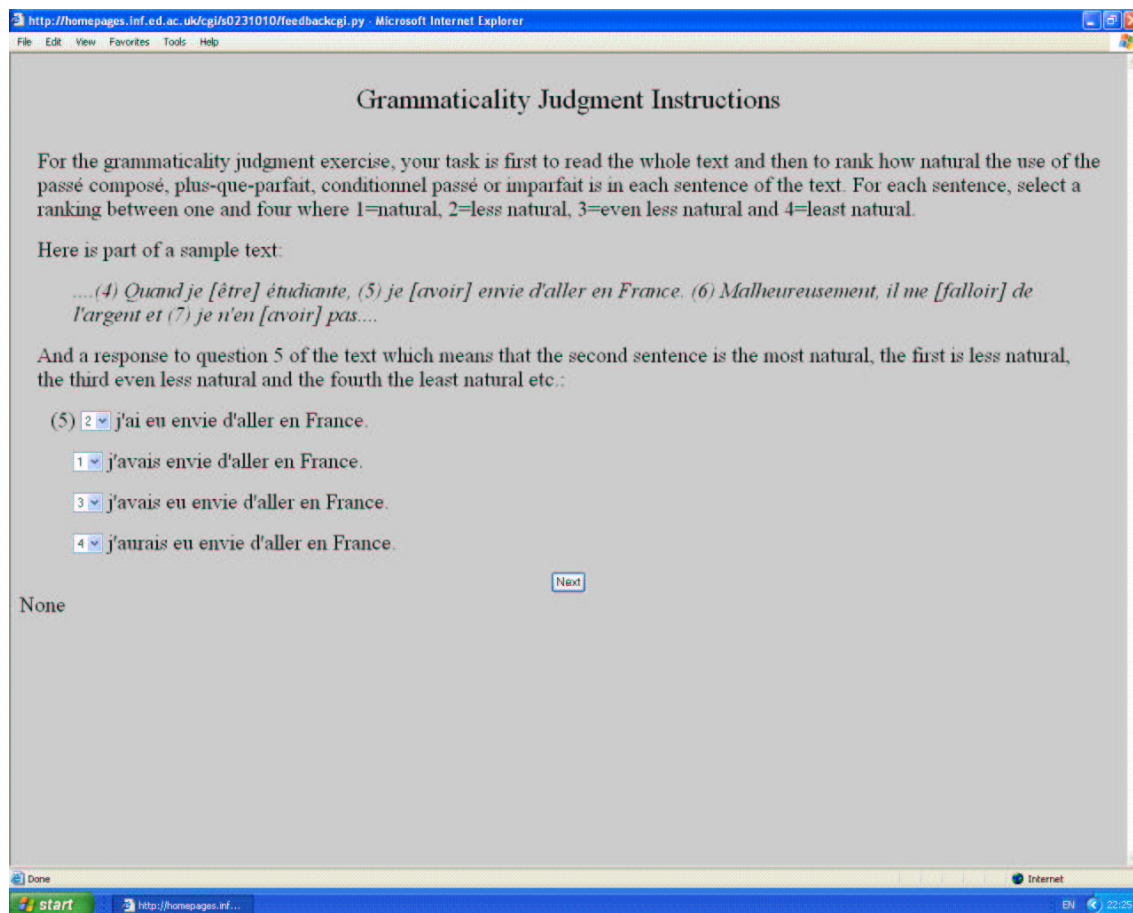
Figure 5.2: Screenshot of the page displaying the instructions for the Grammaticality Judgment Task

The following two pages, illustrated in Figures 5.2 and 5.3 provided detailed instructions on each exercise type as well as a completed sample question. To ensure that there was no confusion, the instructions were also repeated on each page of both tasks (see Figures 5.4 and 5.6). While it may have been advisable for the subjects to complete a sample question before beginning the experimental treatment, it was decided that this was not necessary due to the fact that the tasks were relatively simple and the researcher would be present to discretely check that the subjects were doing the tasks as instructed. This assumption was confirmed, as none of the subjects had difficulty in carrying out the tasks as they were intended.

As explained in Section 5.6.2, the nature of the grammaticality judgment task was such that the analysis of the student input by the system was fairly trivial and the scope for anomalous student input was quite restricted. Therefore,
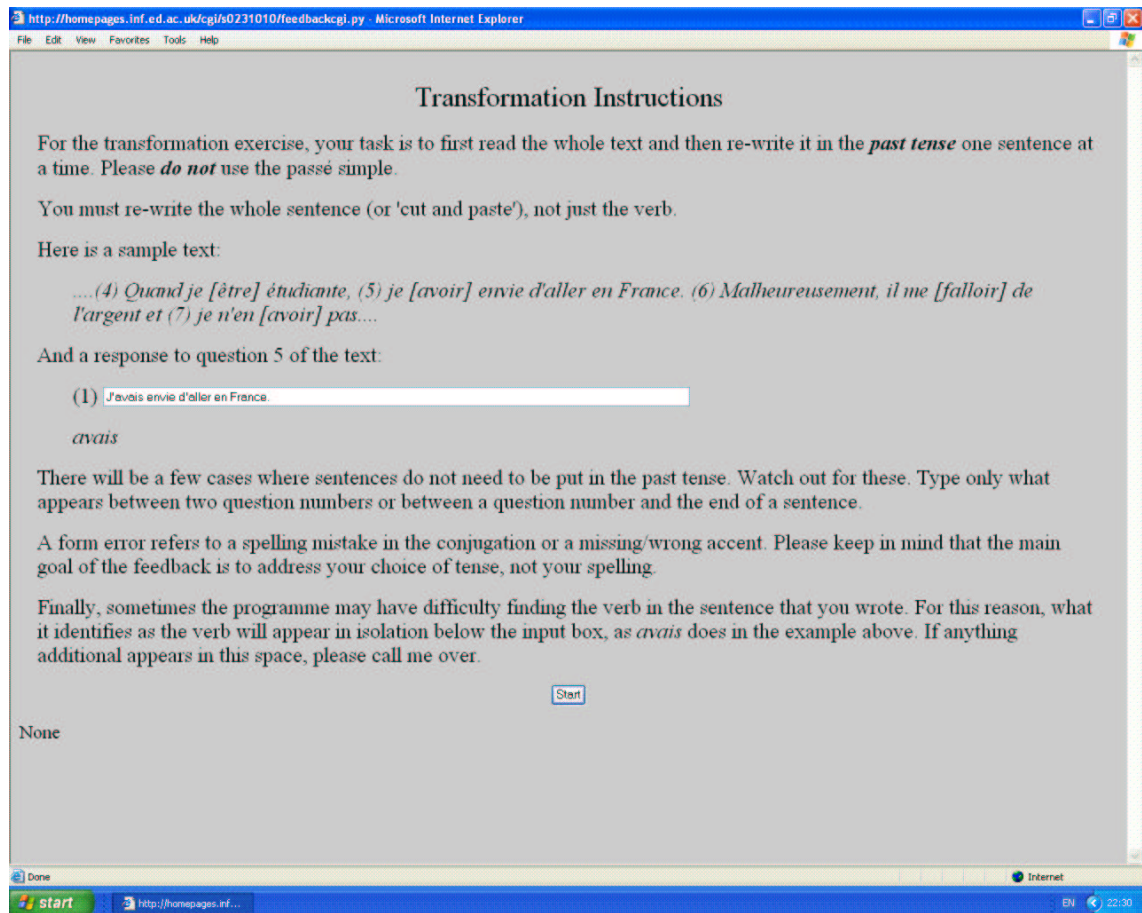
Figure 5.3: Screenshot of the page displaying the instructions for the Transformation Task

the instructions consisted of simply an explanation of the task and a completed sample question (see Figure 5.2). The transformation task, on the other hand, required some additional remarks in an attempt to reduce foreseeable problems to do with the analysis of the input and user considerations (see Figure 5.3). As discussed in Section 5.6.2, the system used a character matching algorithm to identify the verb in the student response. This method of analysis meant that typographic errors elsewhere in the sentence or anomalous whitespace could result in a misidentification of the verb and consequent analysis and feedback generation error. Since this would clearly undermine the user's confidence in the validity of the feedback, it was decided that the verb extracted from the student input should be displayed below the input box so that it would be obvious whether or not the extraction had been successful. If it had, the user could be

confident that the feedback message was reliable, whereas if it had not, the failure could be attributed to a misidentification of the verb, not to the inaccuracy of the feedback.

Notably, the system never failed to correctly identify the verb at any point during the experiment. This was probably largely due to the fact that the subjects cut and pasted each sentence. The instructions included a note that this was permissible for several reasons. Firstly, it meant that the subjects would not be put off by the tedium of re-typing sentences; secondly, it had the benefit of significantly increasing the probability that the analysis of the student response by the system would succeed and thirdly, there was no way of policing the subjects into typing out each word.

Tasks

Having read the instructions, the subjects proceeded directly to the exercises, at which point they were assigned to a feedback group and an exercise combination. Due to the small number of participants, an automatic rotation-style assignment was used according to which the first subject was put into the EI group, the second into the IP group, the third into the Control group and so on. Of the total number of subjects, 11 received EI feedback, 9 IP feedback and 9 no feedback. As discussed in Section 5.6.3, the questions for the exercises were derived from 4 different texts (2 narratives and 2 interviews). Each text comprised one set of questions. These were divided equally across the two exercise types so that each one included both a narrative and an interview. The order in which the sets of questions was displayed was divided into four combinations. Each combination began with a different pair of text and exercise type and alternated between exercise type. Varying the orders of the exercise and text types in this way reduced the possibility of task effects on the subjects' ability to comprehend or produce aspect. As for the assignment of the feedback conditions, the first subject was automatically assigned to exercise combination 1, the second to exercise combination 2 and so forth.

For the grammaticality judgment task, the subjects were to rank the use of the *passé composé*, *plus-que-parfait*, *conditionnel passé* and *imparfait* sentence by sentence according to how natural it was in the context. A version of every sentence
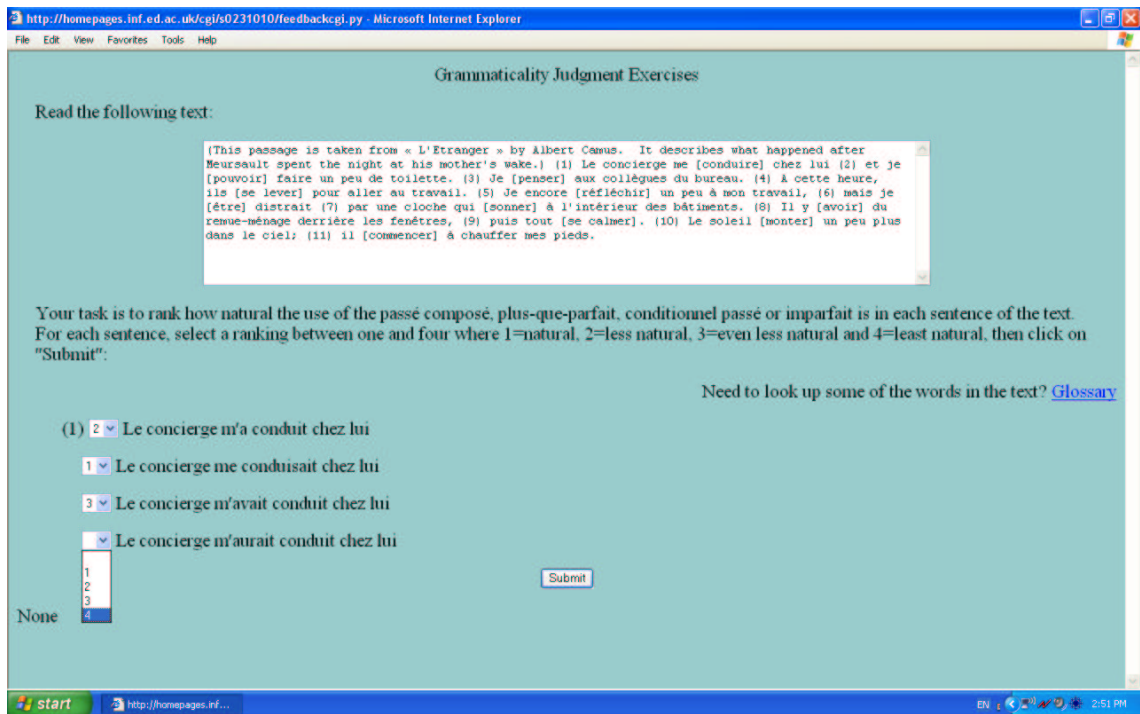
Figure 5.4: Sample Question of Grammaticality Judgment Task

of the text was provided in each of the four tenses and the student marked them on a scale of one to four from natural to least natural, as shown in Figure 5.4. If the subjects gave two sentences the same ranking, they would receive an error message which read, *"You have used the same number twice. Please go back and rank each sentence with a different number."* Likewise, if they failed to rank one sentence, they would receive a message alerting them to the offending sentence, such as, *"You have not ranked the fourth sentence. Please go back and enter a number."*

In the sample question in Figure 5.5, the *imparfait* has been incorrectly marked as the most natural tense according to the context. In reply, the system has offered EI feedback which explains the meaning of the sentence in the *imparfait* and states that it does not make sense according to the context. By clicking on the *Please Try Again* button, the rankings are reset to nothing for the subject to make a second attempt. If a second mistake is made, the EI feedback message appropriate to the error is displayed and a *Next* button appears under the message for the system to advance to the next question. If, at any time, a correct answer is supplied, the system provides a pre-stored congratulatory message of *Bravo!, Good job!, Excellent!, Good!, Well done!, Great!* or *That's right!*.

126
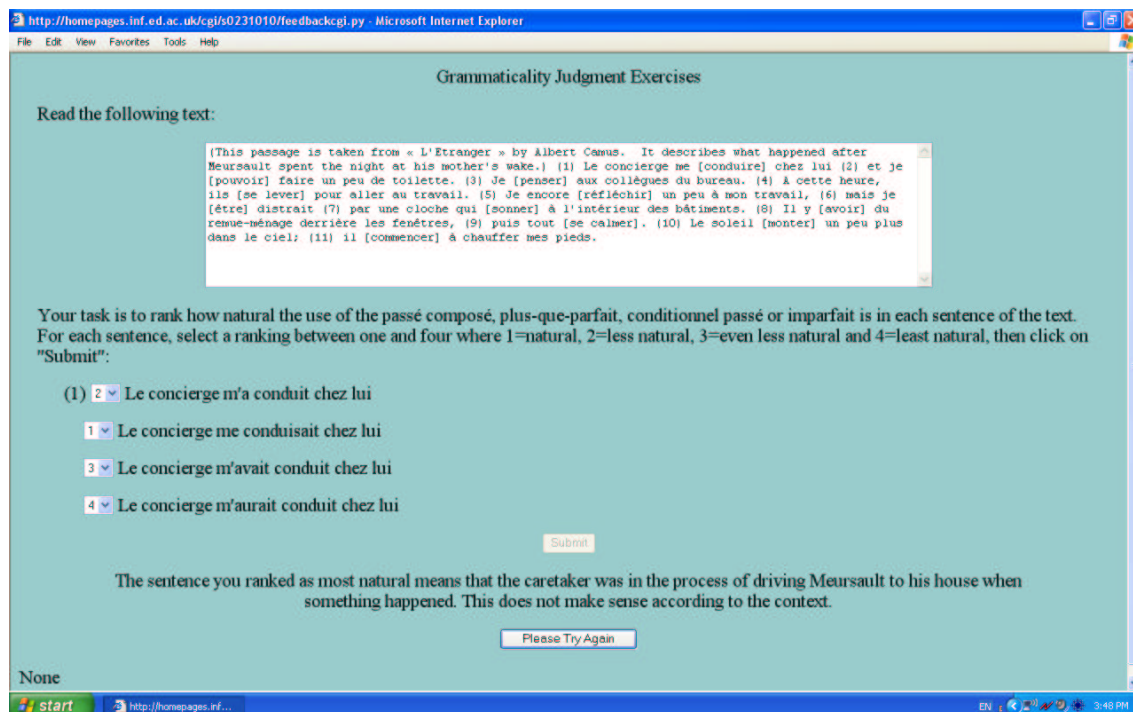
Figure 5.5: Sample Question of Grammaticality Judgment Task with Explicit Inductive Feedback

For the transformation task, a text was given in the present tense and the subjects were asked to re-write it sentence by sentence in the past tense[17], as illustrated in Figure 5.6. The system was designed to cope, to a certain extent, with foreseeable anomalous user behaviour that may have resulted in needless loss of data. Specifically, if the subjects were to accidentally click the *Check Answer* button without entering a response for a question, they would receive an error message which read, *"You have not entered anything. Go back and enter your response."*. Similarly, if they mistakenly used a verb other than the one appropriate for the current question or answered another question altogether, they would be told so and invited to try again.

The system was also capable of managing a number of possible linguistic peculiarities in the student input. Ambiguous input in the form of a lone past participle, which could be interpreted as an attempt at the *passé composé* or the *imparfait*, was addressed with the feedback message, *"Your response is ambiguous.*

---

[17]Instances of the present tense which could have been transformed into a present perfect were excluded from the texts as they would have introduced an additional level of complication for the learners which was beyond the scope of the experiment. Sentences in the passive voice were also excluded on the same grounds, although one was mistakenly included in the comprehension portion of Test A, question 11.
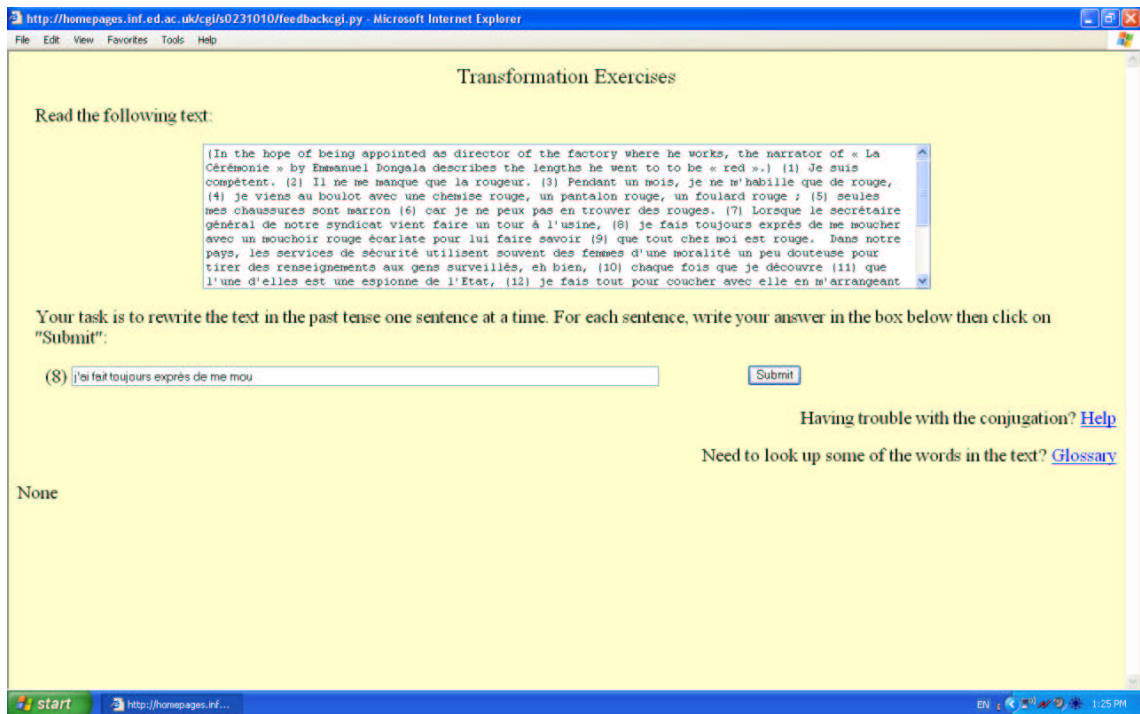
Figure 5.6: Sample Question of Transformation Task

Did you mean to use the imparfait or another past tense? Please try again.". Likewise, if the subjects entered a verb that was a valid simple tense, but not a past tense, they would be asked to try the question again using the *imparfait*, *passé composé*, *plus-que-parfait* or *conditionnel passé*.[18] The rationale behind this was to be able to handle input in the present tense. This was desirable because it was quite likely that the present tense would be used instead of the past either because it was an acceptable choice in some contexts[19]or because the subjects copied and pasted a sentence without modifying the tense. Lastly, input which was target-like in tense but not in form was acknowledged by the system with one of two error messages, either *"Good! (except you've made a form error. We'll move on to the next question anyway...)"* or *"Good! (except you've made an agreement error or used the wrong auxiliary verb. We'll move on to the next question anyway...)"*. The phrasing of these messages sought to alert subjects to their form errors but to keep them from over-focusing on them.

---

[18]Note that this was not the case for composed tenses. If a valid composed tense was entered which was not a past tense (e.g. a well-formed future anterior), it would be analyzed as an ill-formed past tense.

[19]The present conditional was a possible alternative as well, but only in one case.
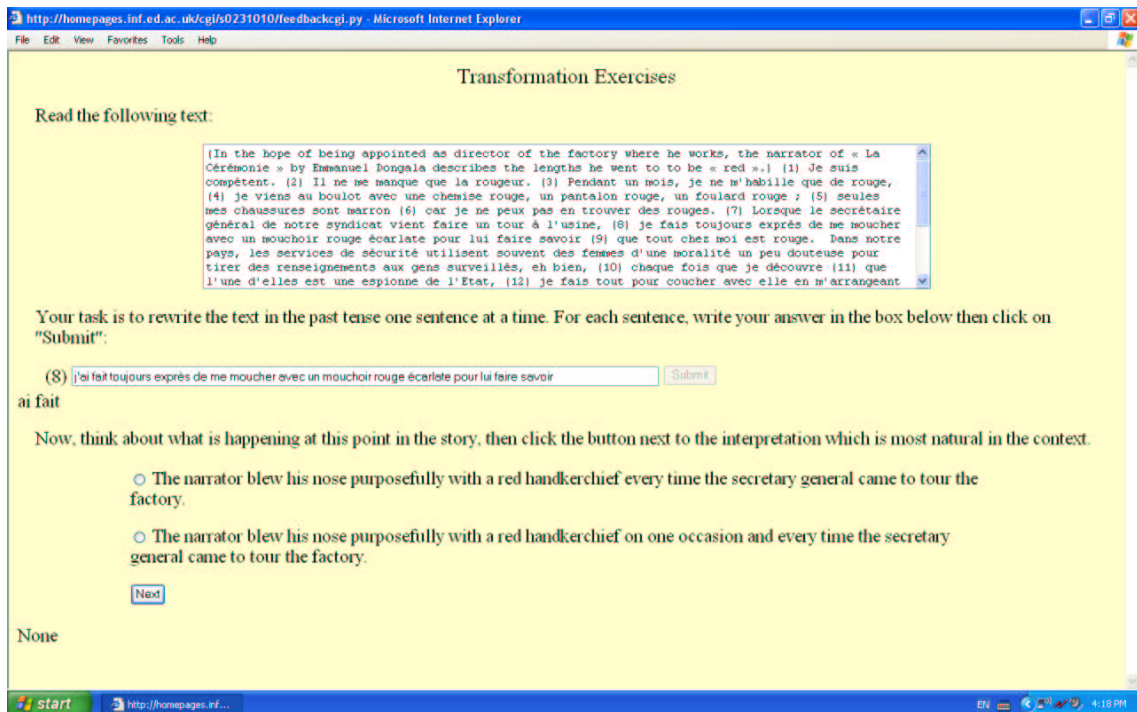
Figure 5.7: Sample Question of Transformation Task with Input Processing Feedback (Part 1)

Figure 5.7 shows a sample question in which the *passé composé* has been used in a context where the *imparfait* is appropriate and the system has generated the first part of some IP feedback. The feedback presents an interpretation of the target sentence in the tense of the subject's response (in this example the second one which is a *passé composé*) and in the appropriate tense for the context (in this example the first one which is a *imparfait*). The subject chooses which of the interpretations is more natural according to the context[20], submits the response and is given the second part of the IP feedback which indicates whether the interpretation is the appropriate one or not. In the sample question, the subject has chosen the right interpretation, as shown in Figure 5.8. Had the wrong interpretation been chosen, the message would have read, *"No, the other interpretation is more natural in the context of the story."*.

---

[20]Note that the display of the interpretations was randomized so that the appropriate interpretation did not always appear in the same place.
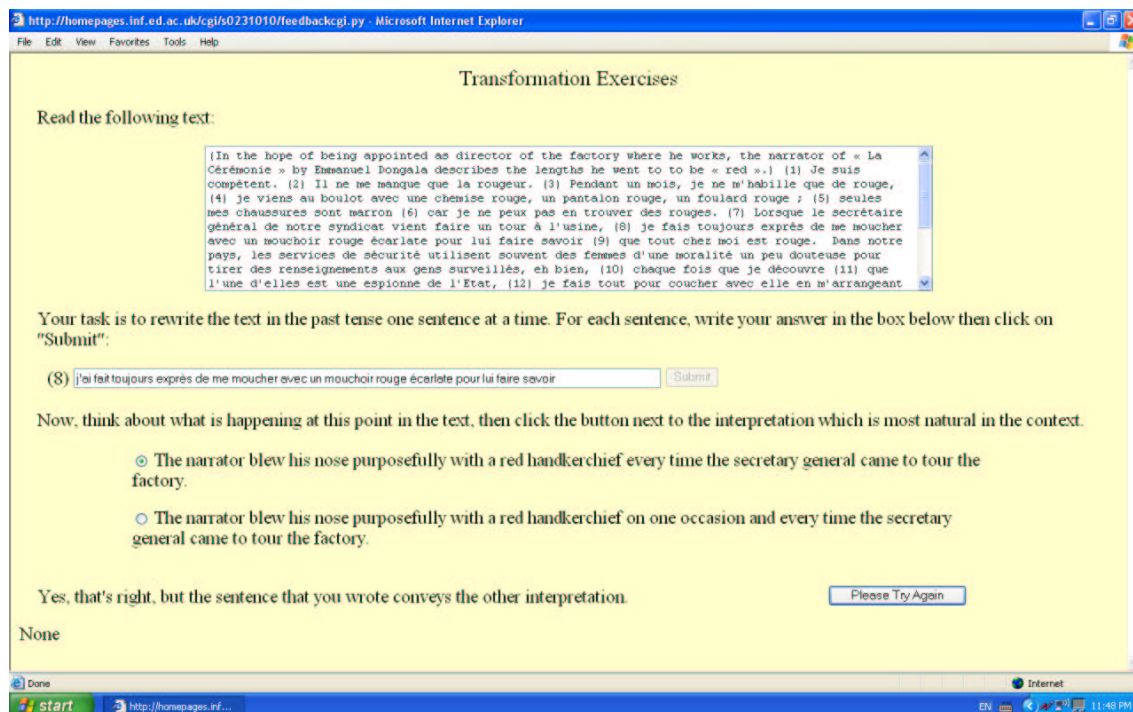
Figure 5.8: Sample Question of Transformation Task with Input Processing Feedback (Part 2)

5.8.3   Post-Testing

Post-testing occurred immediately after and three weeks following the experimental treatment, using test A, B or C, according to the subject's assigned combination order. The results of the first post-test were used to measure the immediate effects of the experimental treatment while the other sought to determine whether these effects had been retained over time.

5.9   Analysis and Results

In order to establish a model against which the learners' aspectual competence could be measured, native speakers of French were needed to complete the experimental materials.

5.9.1   Establishing Native-Speaker Norms

The native-speaker subjects were recruited from the French Section of the Division of European Languages and Cultures and via an advertisement on the

Careers Service website at the University of Edinburgh as well as by word of mouth. There were 10 in total, 7 female and 3 male. While information on the subjects' exact ages was not collected, they were all in their early twenties, except for one who was in her late fifties. Eight of the ten native subjects were ERASMUS[21] students who were on exchange from France and who had arrived in Edinburgh an average of 5 weeks before participating in the experiment. Of the remaining two, one had been living in a non-francophone country for a year and the other for 19 years. All of the subjects were paid £5, or the equivalent in foreign currency, for their participation.

Native-Speaker Profile Questionnaire

The native-speaker subjects completed a questionnaire to confirm that they were all primarily using French on a daily basis and to take account of any factors that may have influenced their language state (i.e. attrition) or heightened their metalinguistic knowledge (e.g. teaching French). As the recruitment of native-speakers was carried out via the French Department at the University of Edinburgh, most subjects were expected to be living in the Edinburgh area. However, in order to maximize the potential number of native-speaker subjects, francophones residing in a French-speaking area were also targeted. The questionnaire was divided into two parts to account for this difference in linguistic environment: one for native-speakers currently residing in a French-speaking country and the other for those living outwith a French-speaking country. All of the native-speaker subjects belonged to the latter category at the time of the study and thus were asked questions on the following points:

- the approximate date of arrival in the non-francophone country
- the amount of time in hours per day, week or month spent speaking French
- the number and duration of annual visits to a francophone country
- any current French teaching, the approximate date when the teaching commenced and the levels taught
- any regular use of languages other than French and English, the name(s) of the language(s) and the frequency of usage per day, week or month

---

[21]ERASMUS (European Community Action Scheme for the Mobility of University Students) is the European Commission's educational programme for administering student exchanges between universities within member and candidate countries of the European Community as well as countries within the European Economic Area.

The exact questions, followed by the native-speaker responses, can be found in Appendix D.

Native-Speaker Data

To establish a benchmark for measuring learner improvement from pre to post tests, each native-speaker completed tests A, B and C and their responses for each question were tabulated. Some adjustments were required, the details of which are explained below, beginning with the comprehension data.

Scoring the native-speaker comprehension data was straightforward, as there was no opportunity for ambiguous, unexpected or missing answers (The reader is referred back to Section 5.6.4 for the details of what the comprehension task required and to Appendix A for exemplars of the tests.). The results showed a lack of consensus on 4 of the 27 target-structure questions, where consensus was defined as over 60% agreement. Although the questions had been checked by a francophone prior to the experiment, further consultation[22] on these questions revealed that both options that the native-speaker subjects had to choose from could be considered grammatical. To illustrate this point, we will consider the instance of Test A, Question 2. The context is given in Test A, Question 1 which states that Dawn Halfaker lost her arm in combat in Irak. In Question 2, reproduced below, Dawn Halfaker responds to the concerns of parents, posed during radio and television interviews in the United States, about the regularity with which women took part in combat in Irak:

> *Dawn Halfaker a répondu que* ⸺⸺⸺⸺⸺⸺
> Dawn Halfaker replied that ⸺⸺⸺⸺⸺⸺
> *des femmes prenaient ainsi part au combat.*
> women took part in battle in this way.
>
> *(a) sur cette mission particulière*
> (a) on this/that particular mission
> *(b) normalement*
> (b) normally

---

[22]I gratefully acknowledge the help of Alice Foucart, a colleague in Psycholinguistics from the Department of Psychology at the University of Edinburgh, in articulating her intuitions and explaining semantic nuances with regard to aspect as a native-speaker of French.

We expected that the native-speakers would view the mission, referred to in option (a), as being complete and, therefore, not compatible with the imperfect *prenaient part*. However, while the mission could certainly have been finished, there was nothing in the context to confirm this. In other words, it is entirely plausible that at the time Dawn Halfaker was giving these interviews, the mission was still going on in Irak. This view of the mission as being incomplete would license the imperfect tense on the verb, thereby rendering options (a) and (b) as grammatical.

The grammaticality of both options was also confirmed by the same native speaker for a further 4 target-structure questions. In these instances, however, the native-speaker subjects came to a 70% or 80% consensus (in 2 cases each) because, while both options were grammatical, one option was more logical than the other in the context. This was the case, for example, in Test C, Question 8, shown below:

> *Le pape Jean XXIII a supporté avec un stoïcisme admirable une agonie de plusieurs jours en raison d'une tumeur de l'estomac.* ⸻
> Pope John XXIII tolerated several days of agony because of a stomach tumor with admirable stoicism. ⸻
> *il a eu des hémorragies, provoquées par la tumeur. Il est mort, âgé de 81 ans.*
> he experienced hemorrhages caused by the tumor. He died, aged 81.

> *(a) Le 3 juin 1963,*
> (a) On June 3, 1963,
> *(b) Pendant plusieurs jours,*
> (b) For several days,

We see, first of all, that option (b) is grammatical with both perfective and imperfective marking on the verb. In the *imparfait*, *"Pendant plusieurs jours, il **avait** des hémorragies"* ("For several days, he experienced hemorrhages") the duration of the hemorrhages is emphasized, whereas in the *passé composé*, *"Pendant plusieurs jours, il **a eu** des hémorragies"* ("For several days, he experienced hemorrhages") the idea that the hemorrhages were an event with a clear endpoint is highlighted. As both tenses are possible, the choice becomes whether one temporal marker is more logical or appropriate than the other. As hemorrhages tend to occur gradually, it would make more sense to say that they occurred over a period of days,

rather than on a specific date. Since the native speakers came to no clear consensus on these eight comprehension questions, they were excluded from the analysis.

The questions for the production part of the pre and post tests were written in such a way as to prompt the subject to produce not only a particular aspect but also to use a synonymous, if not identical, situation type. In three cases, the natives did not use the expected situation type in their answer. For example, in Question 1 of Test A, the question was supposed to elicit an answer such as "Ils étaient étonnés d'entendre un avion" ("They were surprised to hear a plane") or "Ils ne s'attendaient pas à voir un avion" ("They weren't expecting to hear a plane"). In many cases, however, natives chose to say "Ils ont vu un avion" ("They saw a plane") or "Ils ont entendu un avion" ("They heard a plane") which was not the stative situation type that was expected and was, therefore, discarded from the analysis. There was one further problematic case for which, like the comprehension questions described above, both perfective and imperfective aspect were grammatical.

In total, 12 questions (8 comprehension and 4 production) out of 54 were excluded from the pre and post tests. This meant that there were not enough tokens in each test to conduct fine-grained analyses on the data comparing either production and comprehension or the three target structures. Instead, the analysis considered only the overall test scores.

5.9.2   Learner Data

As stated in Section 5.9.1 above, consensus on aspectual choice among native speakers was defined as over 60% agreement. In fact, there was 100% consensus on only 10 out of the 19 comprehension questions and 13 out of the 23 production questions. To reflect this, the learner tests were scored according to the percentage of consensus that the natives showed on a per question basis. For example, on Test A, Question 1, 90% of the natives produced an *imparfait*, while the remaining 10% supplied the *passé composé*. In this case, a learner who produced an *imparfait* would receive a score of .9 out of 1 and a learner who produced a *passé composé* would be given .1 out of 1.

Scoring Anomalies

Scoring the learner comprehension questions did not pose any difficulty. The only potential problem would have been unanswered questions, but this did not arise because all the tests were checked for completeness when they were submitted.

In addition to the problem of unexpected situation types, encountered as described above in the native speaker data, the production questions were subject to additional anomalies which resulted in a limited number of exclusions of individual questions from the learner data. These included unanswered questions; the use of ambiguous tenses (e.g. "a voulait" which contains both the auxiliary verb for a *passé composé* and the *imparfait* form *"voulait"* and, as such, could be interpreted as an attempt at either tense) and nonsensical answers, where it was clear that the learner had not understood the question.

5.9.3   Preliminary Analyses

Confirming Comparable Level of Difficulty of Improvement Measures (Tests A, B and C)

Before carrying out the main analysis, it was necessary to confirm that the three isomorphic tests, used to measure improvement, were similar in difficulty, as any major discrepancy would have confounded the results. After confirming that the data for tests A ($W(9)=0.96$, n.s.), B ($W(8)=0.86$, n.s.) and C ($W(12)=0.95$, n.s.) were normally distributed using a Shapiro-Wilk test, a one-way independent ANOVA was conducted in which the Independent Variable (IV) was test version and the Dependent Variable (DV) was pre test score. It showed no overall effect, $F(2,26)=.010$, n.s. We concluded from these results that there were no significant differences between test versions and, therefore, tests A, B and C were at a comparable level of difficulty.

Confirming Uniform Competence of Feedback Groups (EI, IP and Control)

As described in Section 5.8.2, the learners were automatically allotted to one of the three feedback conditions according to a rotation-style assignment. The pre test scores, which were normally distributed for the EI ($W(11)=0.95$, *n.s.*),

	Pre Test		Immediate Post Test		Gain
	$\bar{x}$	s	$\bar{x}$	s	
EI group (n=11)	54.88%	7.82	62.34%	7.70	7.46%
IP group (n=9)	51.51%	7.61	65.88%	9.85	14.37%
Control group (n=9)	55.46%	12.34	63.59%	13.32	8.13%

Table 5.3: Pre and Immediate Post Test Means

IP (W(9)=0.88, *n.s.*) and control (W(9)=0.92, *n.s.*) samples, constituted the DV and were compared using a one-way independent ANOVA to confirm a uniform level of competence across feedback groups (the IV). The results revealed no significant difference in performance on the pre test across groups, F(2,26)=0.48, *n.s.*

5.9.4   Main Analysis

Immediate Results

The main analysis sought to test the hypotheses proposed in Section 5.5. They predicted that competence in the production and comprehension of past tense aspect among learners who received either EI or IP feedback would increase, as measured by differences in pre and post test scores, compared to those who received no feedback. Comparing the means of the three experimental groups, as shown in Table 5.3, the IP did better than the EI and control groups. However, the results of a one-way repeated-measures ANOVA, where the IV was feedback type, composed of three levels (EI, IP and control), and the DV was test score (pre vs post), revealed that, while all groups made significant improvement from pre to post test, F(1,26)=16.01, $p < .01$, r=.62, there was no significant interaction effect between test scores and feedback type, F(2,26)=0.76, *n.s.* Contrary to the predictions of the experimental hypotheses, these results indicate that all groups made significant improvement from pre to post test, whether they received feedback or not.

A multiple regression analysis which tested whether feedback was a predictor of the change in scores from pre to immediate post test further confirmed the findings of the ANOVA. Using the control group as a baseline category to which the two feedback groups were compared, the idea was to establish whether as the group changed from control to either of the feedback conditions, there might be a significant increase in gain scores. The results revealed an R^2 value of .056,

	Immediate Post Test		Delayed Post Test		Gain
	$\bar{x}$	s	$\bar{x}$	s	
EI group (n=11)	62.34%	7.70	60.13%	12.67	-2.21%
IP group (n=9)	65.88%	9.85	55.15%	14.84	-10.73%
Control group (n=9)	63.59%	13.32	59.11%	10.15	-4.48%

Table 5.4: Immediate and Delayed Post Test Means

	Pre Test		Delayed Post Test		Gain
	$\bar{x}$	s	$\bar{x}$	s	
EI group (n=11)	54.88%	7.82	60.13%	12.67	5.25%
IP group (n=9)	51.51%	7.61	55.15%	14.84	3.64%
Control group (n=9)	55.46%	12.34	59.11%	10.15	3.65%

Table 5.5: Pre and Delayed Post Test Means

meaning that 5.6% of the variance in gain scores could be explained by the group to which a subject belonged. This was not a significant amount, $F(26,2)= .764$, *n.s.* Thus, a significant correlation between feedback group and gain score could not be established.

Delayed Results

Comparisons of the pre and immediate post test means with those of the delayed post test are presented in Tables 5.4 and 5.5. Since the analysis of the pre and immediate post test results showed that the provision of feedback did not have a significant effect, we did not expect an effect on the delayed post test results. This assumption was confirmed, as the delayed post test scores were significantly lower than the immediate ones, $F(1,26)=4.41$, $p < .046$, $r=.38$, and there were no significant differences between feedback groups, $F(2,26)=.85$, *n.s.* Thus, regardless of the provision of feedback, the learners did not maintain the progress that they had shown on the immediate post test. This was further confirmed by the fact that the delayed post test scores were not significantly higher than the pre test ones, $F(1,26)=2.65$, *n.s.*

5.9.5   Further Analysis

While the learners did not exhibit significant improvement due to feedback on the immediate or delayed post tests, the possibility remained that the feedback

Gain Scores (%)					
EI		IP		Control	
S5	-9.07	S19	-5.58	S7	-25.50
S17	-3.60	S14	0.05	S20	-11.46
S26	-0.78	S29	11.73	S27	1.67
S11	0.96	S6	13.21	S23	10.42
S9	4.88	S2	16.87	S13	12.92
S12	6.00	S8	16.90	S24	14.63
S4	6.23	S16	17.66	S18	19.30
S10	14.27	S28	28.35	S25	25.21
S1	14.34	S3	30.17	S15	25.97
S21	20.45				
S22	28.36				

Table 5.6: Gain Scores from Pre to Immediate Post Test

was having an effect while the learners were carrying out the experimental tasks, but that they were not yet able to generalize the knowledge beyond these tasks. To test this, the number of errors made in the first and last 25% of the experimental questions was tabulated for each learner[23]. Comparison of the scores by group showed that the EI and IP groups were making 7% and 4% fewer errors in the last quarter of the experimental questions as compared to the first quarter. The control group, on the other hand, was making 1% more errors. The means show a pattern in favour of the experimental groups, particularly the IP group. However, neither the difference between groups, $F(26,2)=1.08$, *n.s.*, nor their overall improvement, $F(26,1)=2.44$, *n.s.* were significant.

5.9.6 Individual Results

Looking at the gain scores from pre to immediate post test displayed in Table 5.6, 6 out of the 29 subjects had negative scores. While three of these were in the EI group (S1, S2, S3), their regression was not as dramatic as that of the two belonging to the control group (S1, S2). The control group had an average negative score of -18.48% compared to -4.48% for the EI group. In the IP group, only one learner received a negative score which was relatively low at -5.58%.

---

[23]Only errors concerning the target distinction between the *imparfait* and *passé composé* were included in these scores.

At the opposite end of the spectrum, several learners made spectacular gains of over 20% from pre to immediate post test. Subjects 8 and 9 in the IP group had slightly higher scores, averaging a 29.26% improvement. However, this was comparable to the two learners in each of the EI (S10, S11) and control groups (S8, S9) who showed comparable progress, respectively averaging a 24.41% and 25.59% increase.

Looking at the top 50% of gain scores, the IP and control groups showed an advantage over the EI group. Most learners in the IP group made gains of 16.87% or more and members of the control group progressed 12.92% or more. In the EI group, however, most subjects scored only upwards of 6% better.

5.9.7 Summary of Results

The results of the above analyses are summarized as follows:

- Completing the experimental tasks resulted in significantly improved aspectual competence among the learners overall. The Input Processing group showed a 14.37% gain from pre to immediate post test, as compared to a 7.46% gain by the Explicit Inductive group and an 8.13% gain by the control group. However, the contribution of the feedback to improvement was not substantial enough to be anything other than a chance finding.
- 5.6% of the variance in gain scores between pre and immediate post tests could be explained by the group to which a subject belonged. This amount was not statistically significant.
- After a period of three weeks, the learners did not maintain the improvement exhibited directly following the experiment at a significant level.
- Analysis of performance during the experimental treatment showed that the Explicit Inductive and Input Processing groups were making fewer errors and the control group were making more errors. However, these results were not statistically significant.
- There were learners in all three groups who performed worse on the immediate post test than they had on the pre test. The most egregious scores belonged to members of the control group.

- There were also learners in all three groups who performed over 20% better on the immediate post test. The best scores were attained by members of the Input Processing group.
- Most learners in the Input Processing group improved by 16.87% or more. Similarly, the majority of subjects in the control group progressed 12.92% or more. In the Explicit Inductive group, on the other hand, most learners improved by only 6% or more.

CHAPTER 6

Discussion and Conclusion

6.1 Introduction

Discussion of the experimental findings will be divided into an overall assessment and an assessment by feedback group. The former will consider reasons for the results as well as how they fit with previous research. The latter will include a summary of findings, followed by possible explanations for the different tendencies observed between groups and further explanation of the null results. The chapter will finish with summaries of the work as a whole and of suggestions for future research.

6.2 Interpreting the Experimental Results

6.2.1 Overall Assessment

We proposed in Section 5.2 that in order to learn aspect, Anglophone learners' categories of grammatical aspect and combinatorial rules must be restructured to reflect the way aspect is expressed in French. We also argued in Section 5.2.1 that this restructuring is logically possible, provided that the learner is exposed to a categorical aspectual form-meaning relationship. It seems that this is often not the case in the input and for this reason we hypothesized that it may not be possible to acquire aspect on the basis of positive evidence alone. Accordingly, the purpose of the feedback was to explain the aspectual meanings associated with perfective and imperfective morphology so that the relationship between

form and meaning would be clear and restructuring could proceed, or else, in the categorical instances, be expedited. The results of the analyses showed that the effect of the feedback was not substantial enough to be attributable to anything other than chance (see Section 5.9.4 for the results of the main analysis). We account for the results below.

Developmental Readiness: Learner Level and the Non-Categorical Instances of Aspect

From the second language acquisition literature on aspect discussed in Section 4.2, we know that the development of a temporal-aspectual system among learners is slow (Dietrich et al. 1995). For instance, learners in the ESF project had spent an average of three years in the host country at the last sampling and not all of them had even reached the further development stage where aspect begins to be marked morphologically. Similarly, although the learners in our study were producing aspectual morphology, their competence was far from native-like after an average of 7.3 years of formal instruction. The challenge of acquiring perfective and imperfective grammatical aspect has also been explicitly attested in the second language acquisition literature. Montrul & Slabakova (2003, p. 355), for example, state that "[d]evelopmental data from a variety of L2 learners of Romance languages suggests the perfective-imperfective grammatical contrast is perhaps one of the most difficult areas of grammar to master." Referring specifically to French, Noyau et al. (1995, p. 205) go so far as to say "although French has a grammaticalised aspectual distinction in the past, there is no evidence in the data that even advanced learners acquire it."

Clearly, the difficulty in acquiring aspect is widely acknowledged. However, there has been no serious attempt in the second language acquisition or second language learning literatures to explain *why* this might be so. To our knowledge, only Andrews (1992) addresses the question directly and he simply hypothesizes, about the acquisition of French among Anglophones specifically, that one possible source of difficulty is in the difference between the L1 and L2 aspectual systems. As discussed in Section 5.2.1, we believe that the difficulty of acquiring aspect lies in the fact that the relationship between aspectual form and aspectual meaning is often not categorical. In other words, there are contexts where native speakers will make divergent aspectual choices that represent subtle differences in meaning (we refer the reader back to Example 5.2 for an illustration). Since, in

these instances, consistent aspectual choices are not made across native speakers and semantic nuance is not necessarily retrievable from the context, aspectual meaning in the input will be either inconsistent or incomplete.

The content for our experimental questions was taken directly from authentic texts found in books and magazines. This was considered desirable on the grounds that they provided realistic examples of aspectual expression which learners would have to be able to deal with in order to acquire aspect. As authentic materials, they contained a substantial number of cases of aspectual marking that were not categorical. For instance, as mentioned in Sections 5.2.1 and 5.9.1, there was unanimous agreement about aspectual marking by native speakers in only 60% of the transformation exercises.

It was clear from their behaviour in answering the experimental questions that our learners' fundamental understanding of the contrast was sketchy at best. Consider, for instance, the marking of habitual meaning in the Dongala narrative. In this text, the narrator describes two habitual events. The text in question is reproduced below maintaining the question numbers and bolding the occurrences of habitual meaning:

(6.1)   Lorsque le secrétaire général de notre syndicat, membre du comité central du Parti, (7) **venait** faire un tour à l'usine, (8) je **faisais** toujours exprès de me moucher avec un mouchoir rouge écarlate pour lui faire savoir que tout chez moi (9) était rouge. Dans notre pays, les services de sécurité c'est-à-dire notre CIA ou notre KGB, utilisent souvent des femmes d'une moralité un peu douteuse pour tirer des renseignements aux gens surveillés, eh bien, chaque fois que je (10) **découvrais** que l'une d'elles (11) était une espionne de l'Etat, je (12) **faisais** tout pour coucher avec elle...
*When the secretary general of our union, member of the central committee of the Party, (7) **came** to tour the factory, I always purposefully (8) **blew** my nose with a red handkerchief so that he would know that everything about me (9) was red. In our country, the security service, in other words our CIA or our KGB, often use loose women to get information from the people being spied on, well, every time I (10) **found out** that one of them (11) was a State spy, I (12) **did** everything possible to sleep with her...*

The habituality of both events is signalled by adverbials, namely *toujours* ("always") in the first sentence and *chaque fois* ("every time") in the second sentence. Despite such explicit marking, errors were common among learners in both groups. For these four questions, we calculated the number of non-target-like answers on the first attempt. We found that on average the learners scored 1.6 incorrect out of 4 which meant that they were making errors 40% of the time. Furthermore, the data for the analysis reported in Section 5.9.5 showed that the learners in the IP and EI groups were making errors 23% of the time on the last quarter of the experimental questions which involved the perfective-imperfective aspectual contrast[1] (i.e. questions necessitating the *plus-que-parfait* and *conditionnel passé* are excluded from this figure).

This leads us to the hypothesis that our learners were not at a developmental stage where they were ready to grapple with the full range of aspectual marking in their second language. To our knowledge, there is no work within the Processability Theory (Pienemann 1998, 2005) framework which identifies the stages through which learners progress in acquiring aspect. However, it is reasonable to assume that learners would acquire categorical instances of aspect before non-categorical ones, given the distributional complexities of the latter in the input. We hypothesize that presenting our learners with non-categorical instances of aspect made the task of identifying aspectual form-meaning relationships too difficult. The idea of presenting these instances of aspect was misguided in light of the aspectual competence of our learners and this issue accounts, in part, for the absence of effect for the feedback.

Furthermore, since the feedback focused on conveying aspectual meaning, it reinforced the non-categorical nature of the aspectual choice. Consider the feedback for Example 5.2 from Section 5.2.1. Recall that this example is a reply to the question: *Was it true that Pierre Blaise was completely ignorant of what happened during the occupation of France during the second world war?* The answer given in the actual interview was perfective: *Oui. Il a fallu tout lui expliquer* (Yes. It was necessary to explain everything to him). However, imperfective marking is also

---

[1]This may seem low compared to the 40% tabulated for Example 6.1. However, the last quarter of the exercises included all occurrences of perfective-imperfective aspect, not just the target structures. Recall that the proportion of distractors in the texts chosen for the experimental questions was 53% (see Tables 5.1 and 5.2) and that they should have posed relatively little difficulty to the learners compared to the target structures.

possible in this context: *Oui. Il fallait tout lui expliquer* (Yes. It was necessary to explain everything to him). In response to imperfective marking on the verb, the IP group received the following feedback:

(6.2)   IP feedback for imperfective marking on Example 5.2
Now, think about what is happening at this point in the text, then click the button next to the interpretation which is most natural in the context.
○ It was necessary to explain everything to Pierre Blaise at a particular moment once and for all.
○ It was necessary to explain everything to Pierre Blaise gradually over a long period of time.

The text gives no further information about what explaining the Occupation to Pierre Blaise entailed. There is nothing to indicate that the explanation might have taken a long time, nor that it was dealt with in a punctual fashion before moving on to some other event. In the absence of further details, it is clear, both from the text and from the feedback, that either the imperfective or the perfective meaning fits in this context. In other contexts (see Example 5.1), one aspectual meaning is used to the exclusion of the other which reinforces the idea that perfective and imperfective meanings are contrastive. We believe that presenting a context in which both meanings are possible aggravates the task of establishing the form-meaning relationship of the perfective-imperfective aspectual contrast.

EI feedback is also problematic for this reason. It conveys that the aspectual choice in this question is not categorical by conceding that imperfective marking is possible. The EI feedback for Example 5.2 in response to imperfective marking on the verb is reproduced below:

(6.3)   EI feedback for imperfective marking on Example 5.2
What you wrote means that it was necessary to explain everything to Pierre Blaise gradually over a long period of time. This is possible but not the most natural response according to the context.

The explanation of our results elaborated in Section 6.2.1 above is consistent with previous research on the effect of feedback on the acquisition of aspect, namely that of Ayoun (2001). Recall from Section 2.3.2 that her analysis showed a statistically significant advantage for the recast group over the traditional grammar instruction group, but not over the modelling group. We believe that there are two important differences between our study and hers. Firstly, while our feedback was consistently delivered immediately following each error, Ayoun (2001)'s feedback was presented at different times for each feedback group. The recast group received their feedback immediately following their response. The traditional grammar instruction group, on the other hand, were given an answer key to check their response after having received some instruction and completed the experimental questions. In the modelling condition, the learners received a pre-emptive model for each target sentence and were then asked to answer a related question. We believe that the recast group were at a distinct advantage over the traditional grammar instruction group because they received their feedback on a question by question basis immediately after their response was made. The modelling group did not receive their model sentence after making an attempt themselves but each question was addressed individually which may have made the difference between their results and those of the traditional grammar instruction group. Thus, we believe that the significant improvement for the recast group over the traditional grammar instruction group is an artifact of the experimental design.

The second important difference between Ayoun (2001)'s experiment and ours lies in the breadth of aspectual expression to which the learners were exposed during the experimental treatment. Our experimental questions exposed the learners to the full range of aspectual expression in French through the use of authentic texts. Ayoun (2001) does not disclose her experimental questions. However, she describes them as "a contextualized short passage" (Ayoun 2001, p. 234). She does provide sample posttreatment tasks, however, which she refers to as "contextualized stories". We reproduce one of them below (Ayoun 2001, Appendix B):

(6.4)   Task 1: Translation Task

Last year, my brother moved to California, where he had a girlfriend. Her name was Carla. He wanted to go to UC San Diego with her. They were both art majors. They rented a small apartment near La Jolla. On the weekends, they went to museums or art galleries. But one day, my dad got very sick. It happened very quickly. My mother called my brother in California, and he came back right away. We were all very scared. My dad had the best doctors, but he died a week later. My brother and Carla decided to move back to Colorado.

This passage is a textbook example of prose that has been contrived by a language teacher to contrast instances of the *passé composé* and *imparfait*. Specifically, it uses temporal markers, such as *last year* and *one day*, to cue perfective aspectual marking and *on the weekends* to cue imperfective marking. It also uses foregrounding and backgrounding to distinguish clearly between the events of the story (e.g. *my mother called my brother in California, he came back right away, my brother and Carla decided to move back to Colorado*), which would be marked as perfective in French, and the circumstances surrounding the story, which would be marked as imperfective (e.g. *where he had a girlfriend, her name was Carla, he wanted to go to UC San Diego with her*). It seems reasonable to assume that the terms "contextualized short passages" and "contextualized stories" essentially refer to the same type of exercise. Thus, it appears that the learners in this experiment were exclusively exposed to categorical examples of aspect.

We believe that Ayoun (2001)'s significant results were due to inconsistencies in the elicitation method across groups and to the use of categorical examples of aspect. Therefore, we are justified in maintaining that the inclusion of non-categorical examples of aspect may have been a contributing factor in explaining the null results of our experiment.

6.2.2   Assessment by Group

Explicit Inductive (EI) Group Summary of Overall Findings

The EI group appeared to be performing well while answering the experimental questions. They were making fewer errors overall at the end of the experiment

compared to the IP and control groups; however, their performance was not significantly superior (see Section 5.9.5).

As for their performance on the pre and post tests, the individual negative scores were not quite as dramatic as those in the control group. However, as a group, their gain score mean was on a par with the controls (see Table 5.3), indicating no observable advantage for this type of feedback. Furthermore, looking at how the majority of learners in each group fared from pre to post test (see Table 5.6 and Section 5.9.6), most learners in the EI group were not making gains as substantial as their counterparts in the other two groups.

Input Processing (IP) Group Summary of Overall Findings

The IP group showed the most promising tendencies, both as a group and on an individual basis. Although their improvement from pre to immediate post test did not reach significance with respect to the EI and control groups, they made the most substantial gain (see Table 5.3). As a group, they were also making fewer errors overall at the end of the experimental treatment, albeit not a significant number fewer (see Section 5.9.5). Individually, scores showing the most dramatic improvement belonged to members of the IP group (see Section 5.9.6). Furthermore, the majority of the IP subjects improved by 16.87% or more which was somewhat higher with respect to the control group and quite a lot higher compared to the EI group (see Table 5.6 and Section 5.9.6).

Explicit Inductive (EI) and Input Processing (IP) Group Findings on Target Structures

Qualitative analysis of individual learner behaviour with respect to the feedback they received on the target structures revealed some differences between the feedback groups. Recall that for each question the learner was given two chances to come up with a target-like answer. We observed that five of the eleven subjects in the EI group made a second error even after having received feedback. Specifically, they either persisted with the same wrong answer or selected a different tense-aspect marker that was also considered non-target-like. The frequency with which this occurred ranged from twice to sixteen times per learner in the EI group. By comparison, only three learners in the IP group made a second error after having received feedback. Furthermore, this only occurred once

or twice per learner. Thus, it appears that IP feedback more reliably resulted in an immediate correction on the target structures than EI feedback did.

The Feedback Per Se

To explain the apparent disparity between the EI and IP groups, let us compare both types of feedback in response to the same question, as shown in Figures 6.1 and 6.2[2]. As shown in Figure 6.1, the EI feedback provides the learners with an interpretation of their non-target-like *passé composé* response and tells them explicitly that their response is problematic. In Figure 6.2, the IP feedback provides two interpretations and asks the learner to choose which one is most natural in the context. One interpretation is the same as that given in the EI feedback and denotes the interpretation of the learners' *passé composé* response. The other is an interpretation which is considered target-like in the context of the narrative and corresponds morphologically to the *imparfait*. If they choose the wrong interpretation, they are told that the other interpretation is more natural in the context. If they choose the right interpretation, they are told so but that the sentence they wrote conveys the wrong interpretation.

The purpose of both feedback types was to make the relationship between aspectual form and meaning clear; however, they do so in different ways. EI feedback highlights conflicts between aspectual meaning and aspectual form. For instance, in the example given in Figure 6.1, the learner is presented with the accomplishment [the narrator purposefully blow his nose] which he or she is to mark with imperfective grammatical aspect in order to invoke the habitual interpretation that is appropriate in the context. In the event that an error is made, the learner needs to know that one of the features of imperfective grammatical aspect is that it conveys habitual meaning. EI feedback does this indirectly by telling the learner that the morphology they used (perfective, in this case) does not convey a habitual meaning. The learner is then left to infer that, as the only other aspectual marker, imperfective grammatical aspect must be used to express habituality.

---

[2]The entire sentence translates as *When the secretary general of our union came to tour the factory, I always purposefully blew my nose with a red handkerchief so that he would know that everything about me was red.* Recall that the task was to transform this sentence into the past tense.

This presupposes, firstly, that it is clear to the learners what meaning is logical or appropriate in this context and, secondly, that they are able to infer the target-like verbal morphology. In this example, the learners had to ascertain that habitual meaning was appropriate in this context. They then had to take the step of inferring that if perfective morphology does not convey habituality, then imperfective morphology does. We believe that, due to their level of proficiency with respect to the difficulty of the texts, some of our learners had difficulty assigning an aspectual meaning to the experimental questions and that this precluded successful inferencing with respect to the morphology. EI feedback would not have been helpful in this respect because it only identified inappropriate meanings. As for those learners who did know what meaning was appropriate, we think that while the feedback may have led them to correct isolated sentences, it did not go so far as to lead them to make the necessary inferences. If they had, we should have observed a reduction in the number of errors on each target structure as they progressed through the exercises as a minimum, if not significant gains from pre to post test as well.

IP feedback focuses first on identifying an appropriate aspectual meaning. For the example in Figure 6.2, IP feedback starts by having the learners identify what a habitual interpretation of [the narrator purposefully blow his nose] is. When the learners succeed, they are told that the sentence they wrote conveys the other (in this case, perfective) interpretation. From this information, they are supposed to infer that the habitual interpretation should be marked with imperfective grammatical aspect. As mentioned above, in most cases, the learners in this group made the correction after receiving the feedback. However, like the EI group, they did not appear to be able to generalize to subsequent questions. We interpret this behaviour to mean, first of all, that the IP feedback was helpful in establishing a target-like meaning. We believe that this led to a higher proportion of immediate corrections for the IP group as compared to the EI group. However, like the EI group, there was no observable reduction in the number of errors on the target structures as the learners in this group progressed through the exercises. We conclude that they were not making the inferences necessary to establish the aspectual form-meaning relationship which would lead to restructuring.
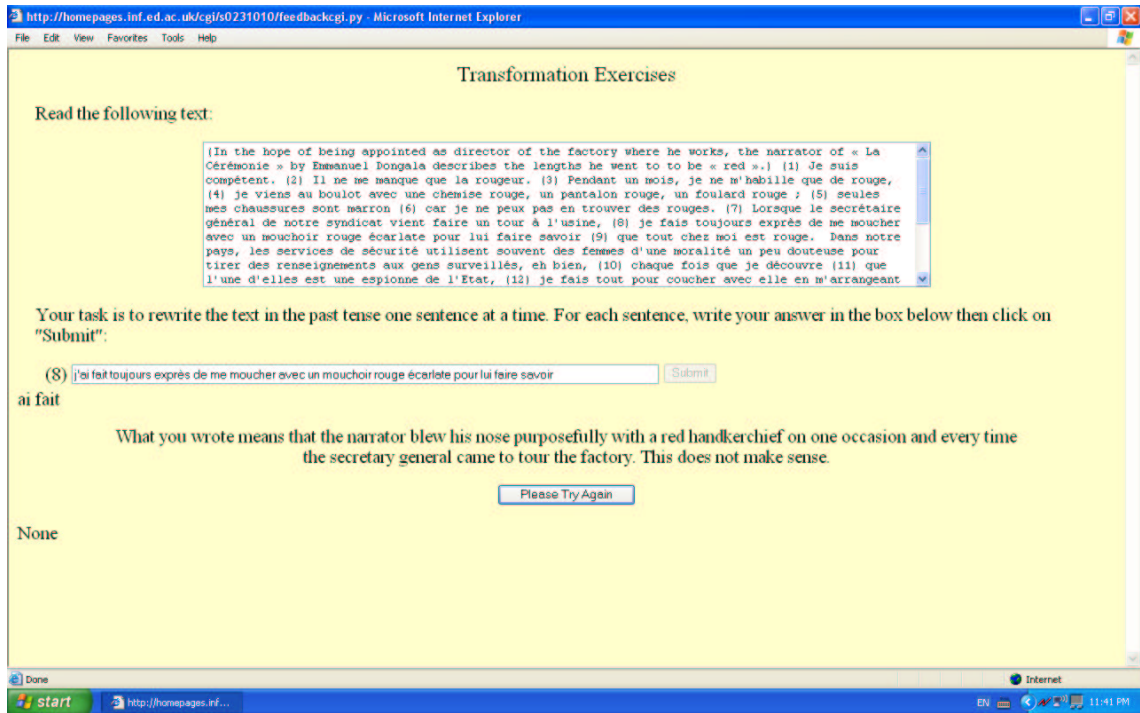
Figure 6.1: Sample EI Feedback in response to Question 8 of Dongala Narrative (Transformation Task)
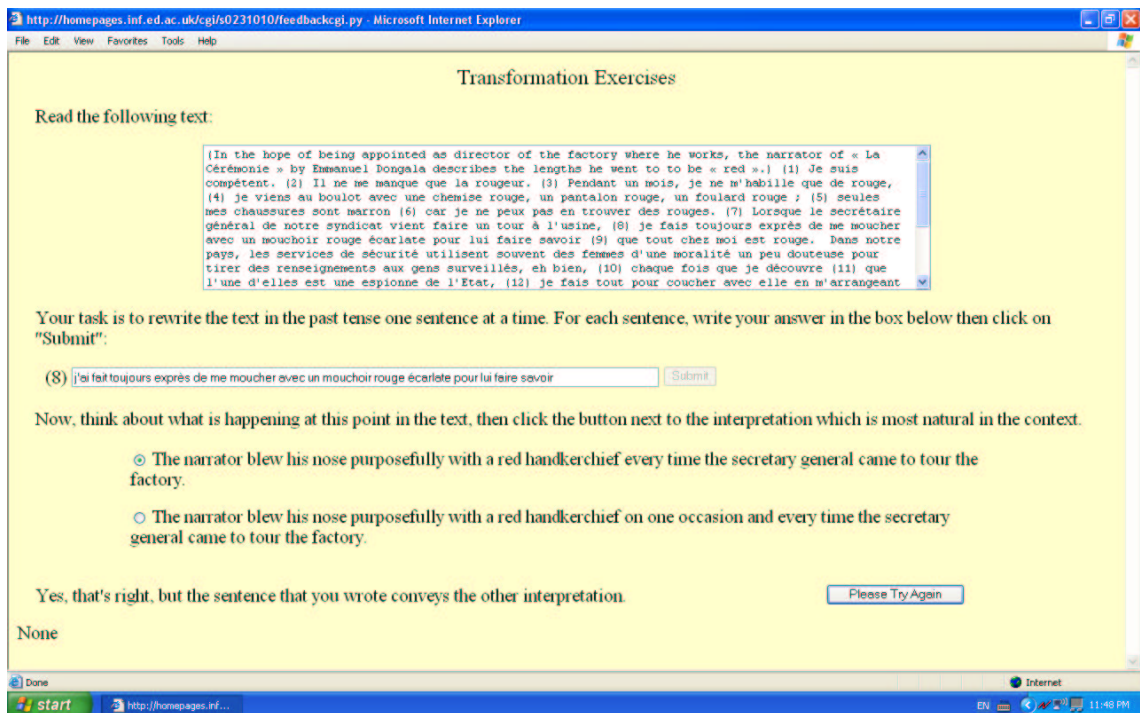
Figure 6.2: Sample IP Feedback (Complete) in response to Question 8 of Dongala Narrative (Transformation Task)

6.3   Chapter Summary

The finding of the experiment as a whole was that the effect of the feedback was not substantial enough to be attributable to anything other than chance. This led us to propose two reasons for which this might have been the case. Firstly, we suggested that our sample of learners were not at a developmental stage where they were ready to deal with a native-like range of aspectual marking. In particular, we proposed that, given the level of our learners, choosing to include instances where aspectual choice was non-categorical was a mistake from the point of view of setting a task that was at an appropriate level of difficulty for learning to have a chance to occur.

Comparing the two experimental groups, we observed different tendencies in the data but, at the same time, identified a unifying theme which served as a further explanation for the null results. We explored their performance overall (i.e. on the pre and post tests as well as on a selection of the experimental questions) and on the target structures during the experimental treatment. Overall, the explicit inductive (EI) group seemed to be responding to the feedback during the experimental treatment but performed similarly to the control group otherwise. The input processing (IP) group, on the other hand, showed improvement across the board. Analysis of the performance of both groups on the target structures showed that while the IP group made more immediate corrections than the EI group, neither group appeared to be able to apply the aspectual information in the feedback to subsequent questions. Our explanation for the disparity in results between feedback groups was that the IP feedback was helpful insofar as it confirmed a target-like aspectual meaning for the learner. However, we also suggested that the learners in neither group were making the necessary inferences about aspectual form which would have led to restructuring.

Referring back to Pinker's terminology, as presented in Section 1.1.3, we conclude that the feedback was useful, in principle. Through the computerization of the task and the feedback, it was possible to identify that the source of the learners' errors was in their combination of aspectual form and aspectual meaning. Furthermore, the feedback was designed to rectify these errors by clarifying the form-meaning relationship. However, due to the difficulty of the target

structures with respect to their level and to their failure to infer the target-like aspectual form, we conclude that the feedback was not used by our learners.

6.4   Conclusion and Future Work

The research reported in this thesis has sought to contribute to the fields of Intelligent Computer-Assisted Language Learning and Second Language Acquisition. From an ICALL perspective, it has seen the quantitative testing of feedback delivered by an ILTS in a field where the majority of past research has been limited to developing and describing prototype systems. To our knowledge, it has also been the first study in ICALL research which has implemented and tested feedback that has been motivated by empirical work on feedback in the second language acquisition literature. This research was also used to motivate and devise new feedback for testing which marked a departure from the current focus of feedback studies on the Interaction Hypothesis and a contribution to the second language acquisition literature on feedback. The work presented here also endeavoured to advance our knowledge of the acquisition of aspect among second language learners. Specifically, while most work has focused on characterizing the process of acquiring the tense-aspect system of a given target language, we gave some much needed attention to considering what effect the external factor of feedback might have on the process. Doing so also allowed us to treat aspect holistically, rather than to limit our investigation to only a single dimension, as work testing the aspect and discourse hypotheses has done. Furthermore, our work identified the intricacies of the learning problem that aspect poses to learners which will be of use for future work on the acquisition of aspect among Anglophone learners of French.

We have found that the provision of Explicit Inductive and Input Processing feedback did not result in statistically significant improvement in the acquisition of past tense aspect in French among Anglophone learners. We attribute the null results to the interaction of several factors. Given the level of our learners, exposing them to a native-like range of aspectual marking (particularly non-categorical instances of aspect) made learning the contrast too difficult. We believe that future research should reserve the full range of aspectual marking for advanced or near-native speakers. Alternatively, for learners at lower levels, exposure

should be limited to instances where the aspectual form-meaning choice is categorical. Furthermore, the learners did not make the inferences about aspectual form which were necessary for the feedback take effect. This could be because of the general difficulty of the task. However, it is also possible that learners, particularly at this level, need feedback which tells them directly what the aspectual form-meaning relationship is supposed to be. Future research could investigate this question.

The results of our experiment and the associated discussion mark the beginning of an interdisciplinary approach to research on feedback which we hope will guide future work in ICALL and second language acquisition.

APPENDIX A

Pre and Post Tests

The pre and post tests are presented in the following pages exactly as the learners saw them. Tests A, B and C of the production questions appear first, followed by Tests A, B and C of the comprehension questions.

A - Part 1

Read each short paragraph and then write a sentence *in the past tense* according to the specifications of each question. Please do **not** use the passé simple. If you are not sure what some of the words mean, ask.

Question 1: Un soir, France et Protée jouent tranquillement un jeu à la maison. Tout d'un coup, ils entendent un avion. Ils ne s'attendent pas à cela. Ils se regardent et sortent de la maison en courant. Ils voient disparaître l'avion derrière la montagne.

Explain France and Protée's surprise:

Ils _____ un avion.

Question 2: A la maison coloniale, il y a des douches séparées : l'une est la douche des boys pour les domestiques et l'autre est celle des propriétaires de la maison. Protée est un des domestiques.

According to this information, state the shower used by Protée in the story:

Protée _____.

Question 3: Pour faire ses trajets *(get around),* France ne se sert pas d'un mode de transport typique, comme un cheval ou un poney. Elle monte à âne *(rides a donkey).*

State France's preferred means of transport:

France _____.

Question 4: Aimée veut de la cuisine française, non pas de la cuisine anglaise que fait le chef de la maison. Elle se fâche contre lui à ce sujet. Puis, sans prévenir, Jonathan Boothby - un ami anglais d'Aimée - rend visite à la famille.

State what Aimée's wishes were for the evening meal in light of Jonathan Boothby's unexpected visit:

Tout d'un coup *(Suddenly),* elle_____que le chef de la maison fasse sa cuisine anglaise.

Question 5: France et Protée sont dans la cabine du générateur. France demande si le générateur brûle. Protée le touche pendant quelques secondes. Peu après, France fait la même chose et retire sa main tout de suite.

State what France's knowledge was about the temperature of the generator as a result of this scene (Please use the verb *savoir* in your answer.):

Dès qu'elle_____le générateur, elle_____

immédiatement qu'il_____très chaud.

Protée et Aimée passent beaucoup de temps ensemble. Ils échangent des regards significatifs mais ils ne parlent jamais ni de leurs sentiments l'un vers l'autre, ni de leur relation.

Explain Protée's and Aimée's behaviour:

Une tension amoureuse_____entre eux.

Question 7: On voit bien qu'il y a des contraintes sociales qui rendent impossible une relation entre Protée et Aimée.

Formulate a hypothesis about their relationship in the absence of these constraints:

Sans ces contraintes,_____, peut-être.

Question 8: Un soir France entend des hyènes *(hyenas)* près de la maison. Elle court à la chambre de sa mère. Armé d'une mitrailleuse, Protée prend garde aux hyènes à la fenêtre de la chambre. France se calme.

Describe France's emotions upon hearing the hyenas:

Soudainement, elle_____des hyènes.

Question 9: Imaginons qu'on est actuellement *(currently)* dans la période de la Révolution culturelle (1966-1976) en Chine. Mao a décrété *(decreed)* que les femmes doivent non seulement porter un costume Mao ou un uniforme de l'armée, mais aussi avoir les cheveux courts et un style unisexe. Les femmes croient que l'uniforme est la plus belle chose au monde et les permanentes et le maquillage sont une horreur.

Give the political reason for which women during the Cultural Revolution had to dress in the way described above:

Parce que_____.

Describe the style for women imposed by Mao's government:

Le gouvernement de Mao_____.

Today fashion is the fourth largest industry in China. Describe women's opinions about fashion at the time of the Cultural Revolution in China:

Les femmes_____que _____.

B - Part 1

Read each short paragraph and then write a sentence *in the past tense* according to the specifications of each question. Please do **not** use the passé simple. If you are not sure what some of the words mean, ask.

****

Here is some useful vocabulary for answering the questions:

rejoindre quelque chose   *to join something (e.g. an organization)*

****

Question 1: La première scène du film se situe dans l'hospice où travaille Lucien. Lucien lave le plancher et il s'arrête un moment pour fusiller *(to shoot)* un oiseau. Peu après, Lucien commence à travailler pour la police allemande.

Identify Lucien's employer previous to the German Police:

   Lucien_____.

Question 2: Quand Lucien arrive chez lui, il rencontre sa mère et son amant. Lucien n'est pas content de voir l'amant. Il lui crie *(shouts)* et puis il part. Il soulage sa colère en fusillant des lapins.

Describe Lucien's emotional reaction to seeing his mother's lover in his house. (Please use the verb *être* plus an adjective in your answer):

   Pendant un moment, il _____.

Question 3: M. Peyssac est le chef du maquis. Un jour Lucien lui rend visite et lui offre deux lapins qu'il a tués. M. Peyssac rejète la demande de Lucien et Lucien finit par travailler pour la police allemande. Cette vie lui donne satisfaction.

State Lucien's **motivation** for visiting M. Peyssac :

   Lucien_____le maquis.

Question 4: Pour avoir à manger, pour s'amuser et pour se procurer un cadeau pour M. Peyssac, Lucien a la même solution.

Describe Lucien's solution in these situations.

   Il _____ des animaux.

Question 5: A un certain moment dans l'histoire, France crie :« J'en ai marre *(fed up with)* d'être juive ! »

State France's desire vis-à-vis being Jewish at that moment, using the verb *vouloir* in your answer:

France n'_____être juive.

Question 6:   Dans le film, France se fâche contre Lucien et même l'insulte. A la fois, elle
            s'amuse bien avec lui et devient son amante.

Describe France's feelings for Lucien :

        Les sentiments de France vers Lucien_____ambigus.

Question 7:   Suivant la révolution féministe en Pakistan, il y a des femmes à Karachi qui
            sont pompistes *(petrol pump attendants)* dans une station-service.
            Cependant, cela reste très étrange dans un pays où le travail d'un pompiste est
            traditionnellement réservé aux hommes.

State the reason for women becoming petrol pump attendants in Pakistan:

        Il y _____.

Question 8:   La fille du propriétaire dirige la station-service. Il n'a pas de fils.

Formulate a hypothesis about the management of the petrol station under different
familial circumstances:

        Si le propriétaire_____des fils, ils

        _____.

Question 9:   Certains hommes pensent que les pompistes n'ont aucune moralité. Alors, ils
            flirtent avec elles.

Describe some men's behaviour towards the petrol pump attendants:

        Ils_____ avec _____ .

Question 10: Un jour, une des pompistes raconte à ses parents le flirt d'un homme. Bien sûr,
            ils se préoccupent immédiatement d'elle.

Describe the parents' reaction to the man's flirtations towards their daughter:

        Ils _____ .

C - Part 1

Read each short paragraph and then write a sentence *in the past tense* according to the specifications of each question. Please do **not** use the passé simple. If you are not sure what some of the words mean, ask.

Question 1: Aux universités américaines, on sépare la contribution des femmes du reste de l'histoire. On choisit entre des cours sur l'histoire des femmes et des cours d'histoire générale qui excluent les femmes.

Describe the choice of courses for previous generations of students wanting to study history in the United States:

Pendant des années, on_____entre suivre un cours sur l'histoire des femmes ou un cours d'histoire générale qui

_____ .

Question 2: Peter Wood, professeur d'histoire à l'université Duke, travaille sur la fusion des cours d'histoire pour qu'il y ait un seul cours qui inclut la contribution des femmes. En faisant ce travail, il redécouvre beaucoup de femmes peu connues, parmi elles Belva Lockwood et Gertrude Ederle:

State the achievements of Belva Lockwood and Gertrude Ederle:

Belva Lockwood, première femme juge à la Cour suprême, (être)_____ candidate à l'élection présidentielle de 1884 ; Gertrude Ederle, première femme à traverser la Manche *(the English Channel)* en 1926,_____de très loin le temps de tous les hommes qui l'avaient précédée.

Question 3: La fin de Jean-Paul II rappelle celle des autres papes du XXe siècle. Mort à 80 ans d'une crise cardiaque, Paul VI se repose à 30 kilomètres de Rome, dans sa résidence d'été de Castelgandolfo parce que ses médecins sont préoccupés de lui éviter les grosses chaleurs *(keeping him out of the extreme heat)*.

State the reason for Paul VI residing at Castelgandolfo at the time of his death:

Ses médecins _____ .

Question 4: Une affiche *(billboard)* montrant un véhicule accidenté, des débris de verre et un tube de rouge à lèvres, accompagnée d'un slogan sur les femmes au volant *(at the wheel)* est le détonateur d'une première campagne antisexiste au Mexique. L'affiche doit être retirée à cause des plaintes *(complaints)* des féministes.

Give the reason for the feminists' complaints about the advertisement:

L'affiche _____qu'elles étaient de mauvaises conductrices.

Question 5: C'est la première fois que des femmes réagissent aux publicités sexistes au Mexique.

Describe women's presumed reactions to sexist advertising ***before*** this incident:

Les femmes _____ .

Question 6: À l'époque du village planétaire *(global village)* et de l'Europe élargie à 25
             Etats, l'Europe revient au sentiment patriotique. II suffit de penser à la réaction
             espagnole face aux terribles attentats du 11 mars ; tous les immeubles
             (buildings) recouverts de drapeaux nationaux espagnols ainsi que le slogan
             récurrent : 'E-spa-ña, E-spa-ña'.

Describe the feelings of Spanish people after the terrorist attacks on March 11:

Du coup *(All of a sudden),* après le 11 mars, les espagnols

_____ .

Question 7: Comme tous les ans, deux Parisiens passent le week-end pascal *(Easter
            weekend)* à faire du ski en Savoie. Samedi en fin de journée, ils prennent un
            remonte-pente *(ski lift)* pour faire une ultime descente avant la nuit. Mais vers
            1800 mètres d'altitude, le remonte-pente s'arrête brusquement. Les deux
            Parisiens savent tout de suite qu'ils seront obligés de passer la nuit sur la piste
            (ski slope). Le lendemain ils sont retrouvés à l'ouverture des pistes en état
            d'hypothermie modérée.

State (still in the past tense) the Parisians' annual Easter weekend holiday:

Ils_____ .

State the Parisians' realization concerning the location of that Saturday night's
accommodation. (Please use the verb *savoir* in your answer):

Ils_____tout de suite qu'ils_____

_____ .

State the ending of the story of the two Parisians trapped on the ski lift:

Ils _____ .

A- Part 2

Complete the following sentences by circling either (a) or (b). If you are not sure what some of the words mean, ask.

1) Dawn Halfaker était femme soldat qui a perdu un bras en combat contre l'ennemie en Irak. De retour aux Etats-Unis, Dawn Halfaker a participé à des émissions *(programmes)* de radio et de télévision. _____, de nombreux parents de jeunes de son âge se sont inquiétés de savoir s'il était normal que des femmes prennent ainsi part au combat.

    (a)   Parmi les téléspectateurs et les auditeurs,
    (b)   Pendant ces émissions,

2) Dawn Halfaker a répondu que_____des femmes prenaient ainsi part au combat.

    (a)   sur cette mission particulière
    (b)   normalement

3) Dawn Halfaker rêvait _____.

    (a)   d'être blessée en combat la nuit avant d'aller en Irak.
    (b)   d'avoir un bras prothétique.

4) Il y avait en Chine une langue inventée il y a plus de quatre cents ans par des femmes, et qu'elles seules pouvaient parler. Cette langue, appelée nushu, permettait aux femmes de parler entre elles de leur intimité et de leurs sentiments profonds, en utilisant un code incompréhensible pour les hommes. _____, la dernière locutrice *(speaker)* de nushu, Yang Huanyi, était née dans le comté *(county)* de Jiangtong.

    (a)   Morte en septembre dernier,
    (b)   Toujours vivante,

5) Yang parlait _____ à ses enfants pour leur transmettre ses connaissances de la langue.

    (a)   régulièrement
    (b)   un jour

6) _____, Yang s'était rapprochée d'un groupe de femmes, `les sept sœurs sous serment'.

    (a)   Après avoir terminé ses études,
    (b)   Avant de se marier,

7) _____,`les sept sœurs sous serment' faisaient autorité dans le domaine de nushu.

    (a)   A l'époque,
    (b)   Pendant quelques années,

8) _____, Yang a vécu avec elles afin d'apprendre toutes les subtilités du nushu.

    (a)   Pendant trois ans,
    (b)   A cette époque,

9)   Le conservatoire de l'héritage culturel chinois permettait à Zhou, un linguiste chinois, d'étudier le nushu _____.

    (a)   jusqu'à ce qu'il ait terminé la rédaction d'un dictionnaire de la langue.
    (b)   autant qu'il voulait.

(10)   _____, Zhou a été bouleversé *(staggered, struck)*. Il n'avait jamais imaginé qu'une telle langue puisse exister.

    (a)   Au moment de la découverte du nushu
    (b)   Chaque fois qu'il y pensait

(11)   Yang Huanyi était invitée à participer à _____.

    (a)   des conférences dans le monde entier.
    (b)   la troisième conférence de l'ONU sur les femmes.

B - Part 2

Complete the following sentences by circling either (a) or (b). If you are not sure what some of the words mean, ask.

(1)   Le nushu a été_____.

   (a)   une langue traditionnelle que le conservatoire de l'héritage culturel chinois voulait préserver.
   (b)   la première langue traditionnelle inscrite sur la liste des langues à préserver par le conservatoire de l'héritage culturel chinois.

(2)   Zhou, un linguiste chinois, s'est lancé dans la rédaction d'un dictionnaire en nushu _____ il s'y intéressait.

   (a)   dès qu' *(as soon as)*
   (b)   parce qu'

(3)   _____Yang étudiait le nushu cinq heures par jour.

   (a)   Pendant trois ans
   (b)   Typiquement

(4)   Zhou était _____.

   (a)   le seul homme à avoir découvert les secrets du nushu.
   (b)   choisi par le conservatoire de l'héritage culturel chinois pour écrire le dictionnaire.

(5)   Lucien a décidé de rejoindre le maquis. Il a demandé la permission à M. Peyssac, chef du maquis mais il lui a dit que non. Lucien a cru qu(e)(')
   _____.

   (a)   il serait amusant de faire partie du maquis.
   (b)   M. Peyssac lui permettrait d'entrer dans le maquis.

(6)   La Gestapo envoyait un petit cercueil *(coffin)* _____.

   (a)   aux familles de leurs prochaines victimes.
   (b)   à la mère de Lucien.

(7)   _____ Lucien n'aurait jamais accepté de travailler pour la police allemande.

   (a)   S'il aurait su les conséquences,
   (b)   Si M. Peyssac lui avait permis de rejoindre le maquis,

(8)   Quand France était petite, elle passait _____.

   (a)   beaucoup de temps à faire des énigmes *(doing riddles)* avec Protée.
   (b)   quelques années en Afrique avec ses parents.

(9) _____, France se sentait mal à l'aise.

    (a)   Après plusieurs mois en Afrique
    (b)   Face à la présence des hyènes un soir

(10) _____, France avait passé son enfance en Afrique.

    (a)   Retournée en Europe en préadolescence,
    (b)   Arrivée comme bébé,

(11) _____, le mari d'Aimée a eu des problèmes de santé.

    (a)   Tout le long du séjour en Afrique
    (b)   A Noël de la première année qu'il était en Afrique

C - Part 2

Complete the following sentences by circling either (a) or (b). If you are not sure what some of the words mean, ask.

Question 1: Passer une nuit entière au-dessus des pistes... C'est pourtant la mésaventure dont ont été victimes _____.

    (a)   ce week-end deux skieurs parisiens.

    (b)   deux skieurs parisiens à cause d'un manque de vigilance du personnel.

Question 2: Vers minuit, les deux skieurs parisiens pensent à sauter du remonte-pente, mais c'est trop dangereux. Ils essaient de crier, mais personne ne les entend. Quelques heures d'après, ils ne sentent plus leurs membres et bougent de plus en plus difficilement. L'un d'eux a cru

    _____.

    (a)   que la seule solution était de se parler pour oublier le froid.

    (b)   un instant mourir sur place.

Question 3: Les deux skieurs parisiens ont commencé à se parler_____ils ne sentaient plus leurs membres.

    (a)   après qu'

    (b)   parce qu'

Question 4: Ils avaient espéré _____.

    (a)   survivre.

    (b)   faire une ultime descente avant la nuit.

Question 5: _____, ils n'auraient pas pris le remonte-pente si tard le soir.

    (a)   Si les skieurs avaient été plus prudents,

    (b)   Si les skieurs auraient su ce qui allait leur arriver,

Question 6: On débattait des moyens à mettre en oeuvre pour aider les femmes à s'adapter au monde de la finance_____sur les femmes. Comme résultat, l'idée de créer une banque nationale féminine est apparue.

    (a)   à plusieurs conférences internationales

    (b)   à une conférence internationale

Question 7: _____, toutes les banques en Venezuela traitaient les femmes d'une façon discriminatoire.

    (a)   Pendant des décennies,

    (b)   Autrefois,

Question 8: Le pape Jean XXIII a supporté avec un stoïcisme admirable une agonie de plusieurs jours en raison d'une tumeur de l'estomac._____ il a eu des hémorragies, provoquées par la tumeur. Il est mort, âgé de 81 ans.

(a) Le 3 juin 1963,
(b) Pendant plusieurs jours,

Question 9: _____surnommait Jean XXIII "le bon pape".

    (a) On
    (b) Un jour, un de ses cardinaux

Question 10: Plus que ses prédécesseurs, Jean Paul II avait une conscience aiguë de la responsabilité historique des grandes religions face aux défis *(challenges)* du monde : la paix, la lutte *(fight)* contre le sous-développement, l'oppression, l'extrémisme religieux._____, il rêvait d'une réconciliation des trois monothéismes issus d'Abraham.

    (a) Pour l'entrée dans le troisième millénaire,
    (b) Pendant toute sa papauté,

Question 11: _____, Jean Paul II avait l'intention de lutter au maximum pour les droits de l'homme *(human rights)*.

    (a) Avant de mourir
    (b) Après être devenu Pape

APPENDIX B

Subject Profile Questionnaire

Note that each question is presented in turn, followed by the response of each subject. Their responses are included exactly as the subjects had entered them into the tutoring system.

Question 1: How old were you when you were first exposed to French?

s1	11
s2	8 or 9
s3	6
s4	11
s5	11
s6	12
s7	5
s8	11
s9	4
s10	8
s11	11
s12	5
s13	11
s14	11
s15	5
s16	10
s17	8
s18	9
s19	11
s20	11
s21	10
s22	10
s23	3
s24	10
s25	4
s26	8
s27	11
s28	11
s29	11

Question 2: How long have you been studying French? (e.g. since Jan 1999)

s1	Since September 1989
s2	since Sep. 1996
s3	since 1993
s4	since Sept 97
s5	since September 1997
s6	since Sept 1997
s7	1995 (to june 2004)
s8	sept 1997 - june 2004
s9	since 1993/4
s10	since Sept 1999
s11	since August 1998 until May 2003
s12	since 1999
s13	since Jan 2001
s14	since 1999
s15	approx. 10-12 years
s16	since Aug 1998
s17	since august 1999
s18	since september 1995
s19	since 1997
s20	since Aug 1998
s21	since september 1997
s22	since Aug 1998
s23	since August 2002
s24	from Sept 1998 to June 2004
s25	since September 1995
s26	since 1998
s27	since Sept 1997
s28	september 1998
s29	since May1998

Question 3: Please give details of all the formal training you have had in French: (e.g. A-level, French 1)

s1	Junior and High school French (6 years); also one IALS course in IALS (upper elementary class)
s2	″ GCSE,AS level,A level″
s3	″ GCSE, A-level french″
s4	″ GCSE, A level, French 1B″
s5	″ GCSE, A level, French 1B″
s6	″ GCSE,AS and A-level″
s7	A-level
s8	″ GCSE, A-level french″
s9	GCSE and A level
s10	″ GCSE,IB higher level″
s11	Standard Grade and Higher French
s12	Advanced Higher
s13	SQA Standard Grade Higher
s14	Higher and advanced higher
s15	″ grade 1-11,College Intermediate″
s16	Higher level
s17	Higher
s18	″ Common Entrance,GCSE,A-Level″
s19	″ GCSE,A-level″
s20	″ Higher French,Adnanced Higher French″
s21	A-level
s22	Standard Grade and Higher French
s23	SQA Higher French
s24	GCSE and AS Level
s25	A-level
s26	″ Gcse,A-level″
s27	″ GCSE,A-level″
s28	″ GCSE,A-level″
s29	A level French

Question 4: Do you ever do anything French-related outside of coursework such as going out with francophone friends or watching French films?
What do you do?
How often do you do this/these things?
(number of times per week?/month?/year?)

s1	no		
s2	yes	"watch french films, listen to french music"	maybe 3-5 times a month.
s3	yes	go out and see french films	twice a year
s4	yes	watch french films	twice a year
s5	yes	Watch French films	Once a month
s6	yes	French films	Every few months
s7	yes	watch films	a couple of times a year
s8	no		
s9	no		
s10	no		
s11	no		
s12	no		
s13	no		
s14	no		
s15	yes	films/television in Canada	once/day
s16	no		
s17	no		
s18	yes	watch french films	a few times a year
s19	yes	Watch french films	about twice a year
s20	yes	Have been to France on different trips/exchanges	Once a year (average)

Question 4: Do you ever do anything French-related outside of coursework such as going out with francophone friends or watching French films?
What do you do?
How often do you do this/these things?
(number of times per week?/month?/year?)

s21	no		
s22	yes	French Films at home and GFT	A few times a year back when I was learning French
s23	yes	watch movies/ French friends	5 times/year
s24	yes	Visit grandparents in France	Once a year
s25	no		
s26	no		
s27	yes	A few films	Only about 5 ever
s28	no		
s29	yes	"Watch French films, read french newspaper"	"About three or four times a month, once a week"

Question 5: Have you ever been to a French-speaking country?

How many times?

For how long? (days?/months?/weeks?)

s1	yes	2	3 days
s2	yes	3	a month each time
s3	yes	5	2 months
s4	yes	6	6 x 2 weeks
s5	yes	6	A few weeks
s6	yes	3	One week each time
s7	yes	15	" a long time,lived in Brussels when 5 and dad lives in France now go for about 2 weeks twice a year"
s8	yes	2	2 weeks
s9	yes	6	total of about 11 weeks
s10	yes	4	3 weeks
s11	yes	1	3 weeks
s12	yes	10	weeks
s13	yes	6	2 weeks
s14	yes	3	2 weeks each time
s15	yes	3	few days
s16	yes	7	approx. 2 weeks
s17	yes	6	1-3 weeks
s18	yes	5	2 weeks
s19	yes	3	one week each time
s20	yes	8	total of 12 weeks (average)
s21	yes	10	1 or 2 weeks
s22	yes	4	2 Weeks normally each time
s23	yes	6	2 weeks on average
s24	yes	7	about 2 weeks each year
s25	yes	15	1 month
s26	yes	probably about 10	2 weeks usually
s27	yes	1	2 days
s28	yes	2	2 weeks
s29	yes	2	2

Question 6: Are you currently studying any language(s) other than French? Which one(s)?

How long have you been studying this/these language(s)? (months?/years?)

s1	no		
s2	no		
s3	yes	russian	one month
s4	no		
s5	yes	Latin	8 years
s6	yes	Japanese	One month
s7	no		
s8	yes	Spanish	5 years
s9	no		
s10	no		
s11	yes	Spanish	2 months
s12	no		
s13	no	Latin	1 month
s14	yes	latin	6 weeks
s15	yes	"Italian,Gaelic"	weeks
s16	yes	italian	less than 1 month
s17	no		
s18	yes	Spanish/Italian	5 years/2 months
s19	no		
s20	no		
s21	no		
s22	no		
s23	yes	Spanish	1 month
s24	no		
s25	no		
s26	yes	Japanese and Chinese	1 month
s27	no		
s28	yes	spanish	7 years
s29	yes	"Japanese,Italian"	"Four years, five months"

Question 7: Have you *ever* studied any language(s) other than French
and those (if any) mentioned in question 6? What language(s)?
How long did you study it/them for? (months?/years?)

s1	yes	Japanese	12 months (3 1st-year uni courses)
s2	yes	German	5 years in total
s3	yes	spanish	6 years
s4	yes	" german,latin"	latin 1yr german 5years
s5	yes	"German, Spanish, Ancient Greek"	"1 year for German, 3 months for Spanish, 2 years for Ancient Greek"
s6	yes	German	2 years
s7	yes	"German, Russian, Latin, Spanish"	" 3 years, 3 years, 6 years and 1 year respectively"
s8	yes	German	2 years
s9	yes	"spanish, german, latin, greek(ancient)"	"spanish 1 yr, german 3 years, latin 7 years, greek 1 year"
s10	yes	"German, Russian"	2 Years
s11	no		
s12	yes	Spanish and German	1 year
s13	no		
s14	no		
s15	yes	"Italian, Gaelic"	"on off - 5 years, 1 year"
s16	no		
s17	yes	Spanish	1 year
s18	yes	Japanese	1 month
s19	yes	German	2 years
s20	no		
s21	yes	German and Malay	"German; 2 years, Malay;6 months"
s22	yes	German	1 Year
s23	yes	Latin	5 years
s24	yes	Spanish	1 year
s25	yes	"German, Latin"	"1 year, 6 years"
s26	yes	German	5 years
s27	no		
s28	yes	"german, italian, japanese	"5 years, 3 years, 1 year"
s29	yes	German	One year

Question 7 (continued): Do you still use this/these language(s) regularly?
How much time do you spend speaking this/these language(s)?
(hours per day?/week?/month?)

s1	no	
s2	no	0
s3	yes	only occasional opportunity to speak
s4	no	0
s5	no	
s6	no	0
s7	no	0
s8	no	
s9	no	
s10	yes	once a week
s11	no	
s12	no	
s13	no	
s14	no	
s15	yes	an hour/ month
s16	no	
s17	no	
s18	no	
s19	no	
s20	no	
s21	no	
s22	no	
s23	yes	only reading 1/m
s24	no	
s25	no	
s26	no	
s27	no	
s28	no	0
s29	no	

Question 8: Have you spoken any language(s) other than English from birth?

Which language(s)?

How much time do you spend speaking this/these language(s)?

(hours per day?/week?/month?)

s1	no		
s2	no		
s3	no		
s4	no		
s5	no		
s6	no		
s7	no		
s8	no	0	
s9	no		
s10	no		
s11	no		
s12	no		
s13	no		
s14	no		
s15	no	n/a	n/a
s16	no		
s17	no		
s18	no		
s19	no		
s20	no	0	
s21	no		
s22	no		
s23	yes	German	2 hours/ month
s24	no		
s25	no		
s26	no		
s27	no	0	
s28	yes	spanish	14 hours per week
s29	no		

Question 9: Have you ever lived in a French-speaking country?
How long did you live there? (weeks?/months?/years?)
How old were you when you arrived in the French-speaking
country? (in years)

s1	no		
s2	no		
s3	no		
s4	no	0	0
s5	no		
s6	no		
s7	yes	1 year	5
s8	no	0	
s9	no		
s10	yes	3 Months	14
s11	no		
s12	no		
s13	no		
s14	no		
s15	no	n/a	n/a
s16	no		
s17	no		
s18	no	0	0
s19	no	0	
s20	no		
s21	no	0	
s22	no		
s23	yes	2	17
s24	no		
s25	no		
s26	no	0	
s27	no	0	
s28	no	0	0
s29	no	0	0

Appendix C

Placement Test

Aside from the distribution of points, the document reproduced in this appendix is identical to what the subjects saw.

Complete the following sentences in the present tense: (1.5 points)

1)   Claude et Philippe ne _____ (vouloir) pas étudier à la bibliothèque.

2)   Tu _____ (mettre) un pull tous les jours.

3)   Donne-moi quelque chose à boire. Je _____ (mourir) de soif.

Transform the following sentences into a question using inversion: (2 points)

4)   Les ministres discutent de l'économie.

   _____ ?

5)   Mme DuRocher compte vendre sa maison.

   _____ ?

Rewrite the following sentences in the passé composé: (3 points)

6)   Elle descend faire les courses.

   _____.

7)   Les employés ont des ennuis avec le directeur.

   _____.

8)   Votre gentillesse est appréciée.

   _____.

Complete the following sentences with the plus-que-parfait or the past conditional: (4 points)

9)   Si tu m' _____ (appeler), je _____ (venir).

Complete the following sentence with the correct preposition. If no preposition is required, mark the space with an X: (1 point)

10)   Nous avons passé l'après-midi _____ visiter les musées.

11)   Le professeur est prêt _____ nous aider.

Complete the following sentences by writing the correct form of the adjective in parentheses: (2 points)

12)   Ils ont trouvé un _____ (beau) appartement à Paris.

13)   Je cherche des chaussures _____ (marron).

Complete the following sentences by writing the French equivalent of the words in parentheses: (2 points)

14)   Les poèmes de Richard sont intéressants, mais _____ (ours) sont

_____ (better).

15)   Nous allons écouter _____ (those) disques et _____ (the ones) d'Odile.

Complete the following sentences with the missing pronoun(s):   (1.5 points)

16)   Il y a du poulet au frigo. Tu peux _____ prendre un peu si tu as faim.

17)   Les enfants adorent le musée scientifique. Je vais _____ emmener.

Complete the following sentence with a relative pronoun: (1 point)

18)   Expliquez-moi _____ il s'agit.

Complete the following sentences with the correct form of either the present indicative or present subjunctive of the verb in parentheses: (2 points)

19)   Je répète l'explication pour que tu me _____ (comprendre).

20)   Il est évident que les problèmes _____ (être) graves.

Appendix D

Native Profile Questionnaire

Question 1: Pendant combien de temps habitez-vous dans ce pays?

Question 1: How long have you lived in this country?

(e.g. depuis jan 1999)

(e.g. since jan 1999)

010101	sept 2005
02121985	depuis sept 2005 erasmus
blanche	1986
calice1	septembre 2005
chachou	depuis le 7 septembre 2005
france	1annee
gargouille	sept 2005
sarah	depuis sept 2005
shurikn08	depuis septembre 2005
s05	depuis septembre 2005 erasmus

Question 2: Combien de temps passez-vous à parler en français?

Question 2: How much time do you spend speaking in French?

(heures par jour?/semaine?/mois?...)

(hours per day?/week?/month?...)

010101	1 heure par jour
02121985	3 heures par jour
blanche	4 heures par jour
calice1	8 heures/jour
chachou	6 heures par jour
france	1heure par jour
gargouille	une heure par jour
sarah	1heure par jour
shurikn08	6h par jour
s05	6h par jour

Question 3: Combien de fois par an rentrez-vous dans un pays francophone?

Question 3: How many times per year do you go back to a francophone country?

010101	1
02121985	0
blanche	1 fois tous les 3/4 ans
calice1	2
chachou	4
france	3
gargouille	une fois
sarah	2
shurikn08	1
s05	3

Question 4: Quand vous rentrez dans un pays francophone,

Question 4: When you go back to a francophone country,

pendant combien de temps y restez-vous d'habitude? (en semaines)

how long do you usually stay? (in weeks)

010101	2
02121985	0
blanche	4
calice1	1
chachou	3
france	1
gargouille	2
sarah	2
shurikn08	1
s05	3

Question 5: Enseignez-vous le français?

Question 5: Do you teach French?

Pendant combien de temps enseignez-vous le français? (e.g. depuis jan 1999)

For how long have you taught French? (e.g. since jan 1999)

Quels niveaux enseignez-vous?

What levels do you teach?

010101	no		
02121985	no		
blanche	yes	1989	tous les niveaux
calice1	no		
chachou	no		
france	no		
gargouille	no		
sarah	no		
shurikn08	no		
s05	no		

Question 6: Parlez-vous régulièrement d'autres langues

Question 6: Do you regularly speak any languages other

que le français et l'anglais?

than French and English?

Quelle(s) langue(s)?

What language(s)?

Combien de temps passez-vous à parler cette langue/ces langues?

How long do you spend speaking this/these language(s)?

(heures par jour?/semaine?/mois?...)

(hours per day?/week?/month?...)

010101	no
02121985	no
blanche	no
calice1	no
chachou	no
france	no
gargouille	no
sarah	no
shurikn08	no
s05	no

References

Allen, L. Q. (2000), 'Form–meaning connections and the French causative: An experiment in Processing Instruction', *Studies in Second Language Acquisition* **22**, 69–84.

Allwright, R. L. (1975), Problems in the study of the language teacher's treatment of learner error, *in* H. C. Dulay, ed., 'On TESOL '75', TESOL, Washington, DC, pp. 96–109.

Andersen, R. W. (1986), El desarollo de la morfología verbal en el español como segundo idioma [The development of verbal morphology in Spanish as a second language], *in* J. M. Meisel, ed., 'Adquisición del lenguaje – Aquisição da linguagem [Acquisition of Language]', Vervuert Verlag, Frankfurt, pp. 115–138.

Andersen, R. W. (1989), 'La adquisición de la morfología verbal [The Acquisition of Verb Morphology]', *Lingüística* **1**, 90–142.

Andersen, R. W. (1991), Developmental sequences: The emergence of aspect marking in second language acquisition, *in* T. Huebner & C. A. Ferguson, eds, 'Crosscurrents in second language acquisition and linguistic theories', Benjamins, Amsterdam, pp. 305–324.

Andersen, R. W. (2002), The dimensions of "Pastness", *in* R. Salaberry & Y. Shirai, eds, 'The L2 Acquisition of Tense-Aspect Morphology', Vol. 27 of *Language Acquisition and Language Disorders*, John Benjamins, Amsterdam/Philadelphia, pp. 79–105.

Andersen, R. W. & Shirai, Y. (1994), 'Discourse motivations for some cognitive acquisition principles', *Studies in Second Language Acquisition* **16**, 133–156.

Andersen, R. W. & Shirai, Y. (1995), 'The Acquisition of Tense-Aspect Morphology: A Prototype Account', *Language* **71**(4), 743–762.

Andersen, R. W. & Shirai, Y. (1996), The primacy of aspect in first and second language acquisition: The pidgin-creole connection, *in* W. C. Ritchie & T. K. Bhatia, eds, 'Handbook of Second Language Acquisition', Academic Press, London/San Diego, pp. 527–570.

Andrews, B. (1992), 'Aspect in Past Tenses in English and French', *International Review of Applied Linguistics in Language Teaching (IRAL)* **30**(4), 281–297.

Ayoun, D. (2001), 'The Role of Negative and Positive Feedback in the Second Language Acquisition of the *Passé Composé* and *Imparfait'*, *The Modern Language Journal* **85**(2), 226–243.

Bach, E. (1986), 'The Algebra of Events', *Linguistics and Philosophy* **9**, 5–16.

Bailin, A. (1995), 'Intelligent Computer-Assisted Language Learning: A Bibliography', *Computers and the Humanities* **29**, 375–387.

Bardovi-Harlig, K. (1992a), 'The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language', *Applied Psycholinguistics* **13**, 253–278.

Bardovi-Harlig, K. (1992b), The telling of a tale: Discourse structure and tense use in learners' narratives, *in* L. F. Bouton & Y. Kachru, eds, 'Pragmatics and language learning', Vol. 3, Division of English as an International Language, University of Illinois, Urbana-Champaign, pp. 144–161.

Bardovi-Harlig, K. (1992c), 'The use of adverbials and natural order in the development of temporal expression', *IRAL* **30**, 299–320.

Bardovi-Harlig, K. (1994), Anecdote or evidence? Evaluating support for hypotheses concerning the development of tense and aspect, *in* E. E. Tarone, S. M. Gass & A. D. Cohen, eds, 'Research methodology in Second Language Acquisition', Erlbaum, Hillsdale, NJ, pp. 41–60.

Bardovi-Harlig, K. (1995), 'A narrative perspective on the development of the tense/aspect system in second language acquisition', *Studies in Second Language Acquisition* **17**, 263–291.

Bardovi-Harlig, K. (1998), 'Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense–aspect morphology', *Studies in Second Language Acquisition* **20**, 471–508.

Bardovi-Harlig, K. (2000), *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use*, Language Learning Monograph Series, Blackwell, Oxford, UK/Malden, MA.

Bardovi-Harlig, K. & Bergström, A. (1996), 'The acquisition of tense and aspect in SLA and FLL: A study of learner narratives in English (SL) and French (FL)', *Canadian Modern Language Review* **52**, 308–330.

Bardovi-Harlig, K. & Bofman, T. (1989), 'Attainment of syntactic and morphological accuracy by advanced language learners', *Studies in Second Language Acquisition* **11**, 17–34.

Bardovi-Harlig, K. & Reynolds, D. W. (1995), 'The role of lexical aspect in the acquisition of tense and aspect', *TESOL Quarterly* **29**, 107–131.

Bayley, R. J. (1991), Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking, PhD thesis, Stanford University, Stanford, CA. Unpublished.

Bayley, R. J. (1994), Interlanguage variation and the quantitative paradigm: Past tense marking in Chinese-English, *in* E. E. Tarone, S. M. Gass & A. D. Cohen, eds, 'Research methodology in Second Language Acquisition', Erlbaum, Hillsdale, NJ, pp. 157–181.

Benati, A. (2001), 'A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense', *Language Teaching Research* **5**, 95–127.

Benati, A. (2004), The Effects of Structured Input Activities and Explicit Information on the Acquisition of the Italian Future Tense, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 207–225.

Benveniste, E. (1966), *Problèmes de Linguistique Générale*, Gallimard, Paris.

Bergström, A. (1995), The expression of past temporal reference by English–speaking learners of French, PhD thesis, Pennsylvania State University, State College, PA. Unpublished.

Bergström, A. (1997), 'L'influence des distinctions aspectuelles sur l'acquisition des temps en français langue étrangère [The influence of aspectual distinctions on the acquisition of tense in French as a foreign language]', *Acquisition et Interaction en Langue Étrangère* **9**, 52–82.

Binnick, R. I. (1991), *Time and the verb: A guide to tense and aspect*, Oxford University Press, New York.

Bley-Vroman, R. (1989), What is the logical problem of foreign language learning?, *in* S. Gass & J. Schachter, eds, 'Linguistic perspectives on second language acquisition', Cambridge University Press, Cambridge, pp. 41–68.

Bley-Vroman, R. & Yoshinaga, N. (1992), 'Broad and narrow constraints on the English dative alternation: Some fundamental differences between native speakers and foreign language learners', *University of Hawai'i Working Papers in ESL* **11**, 157–199.

Cadierno, T. (1992), Explicit instruction in grammar: A comparison of input based and output based instruction in second language acquisition, PhD thesis, University of Illinois at Urbana-Champaign. Unpublished.

Cadierno, T. (1995), 'Formal instruction from a processing perspective: An investigation into the Spanish past tense', *The Modern Language Journal* **79**(2), 179–193.

Cadierno, T. (2000), 'The acquisition of Spanish aspectual distinctions by Danish language learners', *Spanish Applied Linguistics* **4**, 1–53.

Camus, A. (1996), *L'Etranger*, Folio plus, Gallimard, Paris.

Carroll, S. (2001), *Input and Evidence: The raw material of second language acquisition*, Vol. 25 of *Language Acquisition and Language Disorders*, John Benjamins, Amsterdam/Philadelphia.

Carroll, S. (2004), Some Comments on Input Processing and Processing Instruction, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 293–309.

Carroll, S. & Swain, M. (1993), 'Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations', *Studies in Second Language Acquisition* **15**(2), 357–386.

Carroll, S., Swain, M. & Roberge, Y. (1992), 'The role of feedback in adult second language acquisition: Error correction and morphological generalizations', *Applied Psycholinguistics* **13**, 173–198.

Catt, M. & Hirst, G. (1990), 'An intelligent CALI system for grammatical error diagnosis', *Computer Assisted Language Learning* **3**, 3–26.

Chapelle, C. A. (2005), Computer–assisted language learning, *in* E. Hinkel, ed., 'Handbook of Research in Second Language Teaching and Learning', Lawrence Erlbaum Associates, New Jersey, pp. 743–755.

Chaudron, C. (1977), 'A descriptive model of discourse in the corrective treatment of learners' errors', *Language Learning* **27**, 29–46.

Chen, L. & Kurtz, B. L. (1989), XTRA-TE: Using Natural Language Processing Software to Develop an ITS for Language Learning, *in* 'Proceedings of the

4th International Conference on AI and Education', Amsterdam, Netherlands, pp. 54–63.

Cheng, A. (2004), Processing Instruction and Spanish *ser* and *estar*: Forms with Semantic–Aspectual Values, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 119–141.

Cheng, A. C. (2002), 'The effects of processing instruction on the acquisition of *ser* and *estar*', *Hispania* **85**, 308–323.

Chi, M., Siler, S., Jeong, H., T., Y. & Hausmann, R. (2001), 'Learning from Human Tutoring', *Cognitive Science* **25**, 471–533.

Chun, A., Day, R., Chenoweth, A. & Luppescu, S. (1982), 'Errors, interaction, and correction: A study of native-nonnative conversations', *TESOL Quarterly* **16**, 537–547.

Clahsen, H. & Hong, U. (1995), 'Agreement and null subjects in German L2 development: New evidence from reaction-time experiments', *Second Language Research* **11**, 57–87.

Clahsen, H. & Muysken, P. (1989), 'The UG paradox in L2 acquisition', *Second Language Research* **5**, 1–29.

Clahsen, H. & Muysken, P. (1996), 'How adult second language learning differs from child first language development: Commentary on Epstein et al.', *Behavioral and Brain Sciences* **19**, 721–723.

Collentine, J. (1998), 'Cognitive Principles and CALL Grammar Instruction: A Mind-Centered, Input Approach', *CALICO Journal* **15**, 1–18.

Collentine, J. (2002), 'On the acquisition of the subjunctive and authentic processing instruction: A response to Farley', *Hispania* **84**(4), 900–909.

Collins, L. (1997), The development of tense and aspect. Paper presented at the Second Language Research Forum, Michigan State University, East Lansing, MI.

Collins, L. (1999), Achieving new states with a familiar activity: A modern times accomplishment? Paper presented at the Annual Meeting of the American Association of Applied Linguistics, Stamford, CT.

Comajoan, L. (1998), The acquisition of past morphology in Catalan as a foreign language: Interaction of lexical aspect and discourse grounding in the L2 acquisition of past morphology of Catalan. Paper presented at the Second Language Research Forum, University of Hawai'i, Manoa, HI.

Comrie, B. (1976), *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*, Cambridge University Press, Cambridge, UK.

Coppieters, R. (1987), 'Competence Differences between Native and Near-Native Speakers', *Language* **63**(3), 544–573.

Corder, S. P. (1967), 'The significance of learner's errors', *International Review of Applied Linguistics in Language Teaching* **5**(4), 161–170.

Courchene, R. (1980), 'The error analysis hypothesis, the contrastive analysis hypothesis, and the correction of error in the second language classroom', *TESL Talk* **11**(2), 3–13.

Cowan, N. (1995), *Attention and memory: An integrated framework*, Oxford University Press, New York.

Criswell, E., Byrnes, H. & Pfister, G. (1992), Intelligent Automated Strategies of Teaching Foreign Language in Context, *in* M. Swartz & M. Yazdani, eds, 'Intelligent Tutoring Systems for Foreign Language Learning', Vol. F80 of *NATO ASI Series*, Springer Verlag, Berlin Heidelberg, pp. 307–319.

Crookes, G. & Rulon, K. (1988), 'Topic continuation and corrective feedback in native-nonnative conversation', *TESOL Quarterly* **22**, 675–681.

Darbelnet, J. (1977), *Pensée et Structure*, The Scribner French Series, 2nd edn, Charles Scribner's Sons, New York.

Day, R., Chenoweth, N. A., Chun, A. & Luppescu, S. (1984), 'Corrective feedback in native-nonnative discourse', *Language Learning* **34**, 19–46.

de Swart, H. (1998), 'Aspect Shift and Coercion', *Natural Language and Linguistic Theory* **16**(2), 347–385.

Debski, R. (2003), 'Analysis of research in CALL (1980–2000) with a reflection on CALL as an academic discipline', *ReCALL* **15**(2), 177–188.

DeKeyser, R., Salaberry, R., Robinson, P. & Harrington, M. (2002), 'What Gets Processed in Processing Instruction? A Commentary on Bill VanPatten's *Processing Instruction: An Update*', *Language Learning* **52**(4), 805–823.

DeKeyser, R. & Sokalski, K. (1996), 'The differential role of comprehension and production practice', *Language Learning* **46**(4), 613–642.

Dietrich, R., Klein, W. & Noyau, C. (1995), *The Acquisition of Temporality in a Second Language*, John Benjamins, Amsterdam/Philadelphia.

Dongala, E. B. (1996), *Jazz et vin de palme*, Motifs, Le Serpent à Plumes, Paris.

Doughty, C. & Varela, E. (1998), Communicative focus on form, *in* C. Doughty & J. Williams, eds, 'Focus on Form in Classroom Second Language Acquisition', The Cambridge Applied Linguistics Series, Cambridge University Press, Cambridge, chapter 6, pp. 114–138.

Dowty, D. R. (1979), *Word Meaning and Montague Grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, D. Reidel Publishing Company, Dordrecht/Boston.

Dowty, D. R. (1986), 'The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics?', *Linguistics and Philosophy* **9**, 37–61.

Epstein, S., Flynn, S. & Martohardjono, G. (1996), 'Second language acquisition: Theoretical and experimental issues in contemporary research', *Behavioral and Brain Sciences* **19**, 677–714.

Farley, A. P. (2001a), 'The effects of processing instruction and meaning-based output instruction', *Spanish Applied Linguistics* **5**, 57–94.

Farley, A. P. (2001b), 'Authentic processing instruction and the Spanish subjunctive', *Hispania* **84**, 289–299.

Farley, A. P. (2004a), The Relative Effects of Processing Instruction and Meaning–Based Output Instruction, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 143–168.

Farley, A. P. (2004b), Processing Instruction and the Spanish Subjunctive: Is Explicit Information Needed?, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 227–239.

Farrar, M. J. (1990), 'Discourse and the acquisition of grammatical morphemes', *Journal of Child Language* **17**, 607–624.

Ferreira-Cabrera, A. (2003), Feedback Strategies for Second Language Teaching with implications for Intelligent Tutorial Systems, PhD thesis, School of Informatics, University of Edinburgh. Unpublished.

Field, A. (2005), *Discovering Statistics Using SPSS*, 2nd edn, SAGE, London.

Flashner, V. E. (1989), Transfer of aspect in the English oral narratives of native Russian speakers, *in* H. Dechert & M. Raupach, eds, 'Transfer in language production', Ablex, Norwood, NJ, pp. 71–97.

Freed, A. (1979), *The Semantics of English Aspectual Complementation*, Reidel, Dordrecht/Boston.

Fum, D., Giangrandi, P. & Tasso, C. (1988), ET: An Intelligent Tutor for Foreign Language Teaching, *in* 'Proceedings of the International Conference on Intelligent Tutoring Systems', Montreal, pp. 462–468.

Gamper, J. & Knapp, J. (2002), 'A Review of Intelligent CALL Systems', *Computer Assisted Language Learning* **15**(4), 329–342.

Garey, H. (1957), 'Verbal Aspect in French', *Language* **33**(2), 91–110.

Gaskill, W. H. (1980), Correction in native-non native speaker conversation, *in* D. Larsen-Freeman, ed., 'Discourse analysis in second language research', Newbury House, Rowley, MA, pp. 125–137.

Gass, S. (2003), Input and Interaction, *in* C. J. Doughty & M. H. Long, eds, 'The Handbook of Second Language Acquisition', Blackwell Publishing Ltd, Oxford, UK, pp. 224–255.

Gass, S. & Varonis, E. (1985), 'Variation in native speaker speech modification to non-native speakers', *Studies in Second Language Acquisition* **7**, 37–57.

Gass, S. & Varonis, E. (1989), Incorporated repairs in NNS discourse, *in* M. Eisenstein, ed., 'The Dynamic Interlanguage', Plenum Press, New York, pp. 71–86.

Gass, S. & Varonis, E. (1994), 'Input, interaction, and second language production', *Studies in Second Language Acquisition* **16**(3), 283–302.

Giacalone Ramat, A. (1995), 'L'expression de l'aspect progressif en italien seconde langue et le rôle des propriétés sémantiques des verbes [The expression of progressive aspect in Italian as a second language and the role of semantic properties of verbs]', *Acquisition et interaction en langue étrangère* **5**, 47–78.

Giacalone Ramat, A. (1997), Progressive periphrases, markedness, and second–language data, *in* S. Eliasson & E. Jahr, eds, 'Language and its Ecology', Mouton, Berlin, pp. 261–285.

Giacalone Ramat, A. & Banfi, E. (1990), 'The acquisition of temporality: A second language perspective', *Folia Linguistica* **24**(3–4), 405–428.

Gregg, K. R. (2003), SLA Theory: Construction and Assessment, *in* C. J. Doughty & M. H. Long, eds, 'The Handbook of Second Language Acquisition', Blackwell Publishing Ltd, Oxford, UK, pp. 831–865.

Grévisse, M. (1949), *Le Bon Usage: Grammaire Française*, Duculot, Gembloux, Belgium.

Guillaume, G. (1929), *Temps et verbe: Théorie des aspects, des modes et des temps [Tense and Verb: Theory of Aspect, Mood and Tense]*, Collection linguistique

publiée par la Société de linguistique de Paris, 27, Librairie Ancienne Honoré Champion, Paris.

Han, Z. (2002), 'A Study of the Impact of Recasts on Tense Consistency in L2 Output', *TESOL Quarterly* **36**(4), 543–572.

Hasbún, L. (1995), The role of lexical aspect in the acquisition of tense and grammatical aspect in Spanish as a foreign language, PhD thesis, Indiana University, Bloomington, Indiana. Unpublished.

Hawkins, R. (1987), 'Markedness and the acquisition of the English dative alternation by L2 speakers', *Second Language Research* **3**, 20–55.

Heift, T. (2001), 'Intelligent language tutoring systems for grammar practice', *Zeitschrift für Interkulturellen Fremdsprachenunterricht (Online)* **6**(2), 1–15. http://zif.spz.tu–darmstadt.de/jg–06–2/beitrag/heift2.htm.

Hilles, S. (1991), Access to universal grammar in second language acquisition, *in* L. Eubank, ed., 'Point counterpoint: Universal Grammar in the Second Language', John Benjamins, Amsterdam, pp. 305–338.

Hoepelman, J. & Rohrer, C. (1980), On the Mass-Count Distinction and the French *Imparfait* and *Passé Simple*, *in* C. Rohrer, ed., 'Time, Tense, and Quantifiers', Niemeyer, Tübingen.

Holland, V. M., Kaplan, J. D. & Sabol, M. A. (1999), 'Preliminary Tests of Language Learning in a Speech-Interactive Graphics Microworld', *CALICO Journal* **16**(3), 339–359.

Holley, F. M. & King, J. K. (1971), 'Imitation and correction in foreign language learning', *The Modern Language Journal* **55**, 494–498.

Hopper, P. J. (1979), Aspect and foregrounding in discourse, *in* T. Givón, ed., 'Syntax and semantics: Discourse and syntax', Academic Press, New York, pp. 213–241.

Housen, A. (1993), L2 acquisition of verb morphology: A case study, *in* B. Kettemann & W. Wieden, eds, 'Current Issues in European Second Language Acquisition Research', Narr, Tübingen, pp. 195–212.

Housen, A. (1994), Tense and aspect in second language acquisition: The Dutch interlanguage of a native speaker of English, *in* C. Vet & C. Vetters, eds, 'Tense and Aspect in Discourse', Mouton de Gruyter, Berlin/New York, pp. 257–291.

Housen, A. (1998), Facteurs semantico-conceptuels et discursivo-fonctionnels dans le développement des systèmes temporo-aspectuels: Aperçu de l'acquisition de l'anglais comme langue étrangère, *in* S. Vogeleer, A. Borillo,

C. Vetters & M. Vuillaume, eds, 'Temps et discours', Peeters, Leuven, Belgium, pp. 257–279.

Hulstijn, J. (1997), 'Second Language Acquisition Research in the Laboratory: Possibilities and Limitations', *Studies in Second Language Acquisition* **19**, 131–143.

Imbs, P. (1960), *L'Emploi des Temps Verbaux en Français Moderne*, Klincksieck, Paris.

Inagaki, S. (1997), 'Japanese and Chinese Learners' Acquisition of the Narrow-Range Rules for the Dative Alternation in English', *Language Learning* **47**(4), 637–669.

Inagaki, S. & Long, M. H. (1999), The effects of implicit negative feedback on the acquisition of Japanese as a second language, *in* K. Kanno, ed., 'The acquisition of Japanese as a second language', John Benjamins, Amsterdam, pp. 9–30.

Ishida, M. (2004), 'Effects of Recasts on the Acquisition of the Aspectual Form *-te i-(ru)* by Learners of Japanese as a Foreign Language', *Language Learning* **54**(2), 311–394.

Iwashita, N. (2003), 'Negative Feedback and Positive Evidence in Task-Based Interaction: Differential Effects on L2 Development', *Studies in Second Language Acquisition* **25**, 1–36.

Jespersen, O. (1931), *A Modern English Grammar on Historical Principles*, George Allen and Unwin, London.

Joos, M. (1964), *The English Verb: Form and Meaning*, The University of Wisconsin Press, Madison, Wisconsin.

Jung, U. O. H. (2005), 'An international bibliography of computer–assisted language learning: Sixth instalment', *System* **33**, 135–185. Available online at http://www.sciencedirect.com.

Kaplan, J. & Holland, V. M. (1995), Application of Learning Principles to the Design of a Second Language Tutor, *in* J. Kaplan, M. Sams & V. M. Holland, eds, 'Intelligent Language Tutors: Theory Shaping Technology', Lawrence Erlbaum Associates, New Jersey, pp. 273–288.

Kaplan, M. A. (1987), Developmental patterns of past tense acquisition among foreign language learners of French, *in* T. R. Dvorak, B. VanPatten & J. F. Lee, eds, 'Foreign language learning: A research perspective', Newbury House, Cambridge, MA, pp. 52–60.

Kihlstedt, M. (2002), Reference to past events in dialogue: The acquisition of tense and aspect by advanced learners of French, *in* R. Salaberry & Y. Shirai,

eds, 'The L2 Acquisition of Tense–Aspect Morphology', Language Acquisition and Language Disorders, John Benjamins, Amsterdam/Philadelphia, pp. 323–361.

Klein, W. (1995), The Acquisition of English, *in* R. Dietrich, W. Klein & C. Noyau, eds, 'The Acquisition of Temporality in a Second Language', Vol. 7 of *Studies in Bilingualism*, John Benjamins, Amsterdam/Philadelphia, pp. 31–70.

Klein, W., Dietrich, R. & Noyau, C. (1995), Conclusions, *in* R. Dietrich, W. Klein & C. Noyau, eds, 'The Acquisition of Temporality in a Second Language', Vol. 7 of *Studies in Bilingualism*, John Benjamins, Amsterdam/Philadelphia, pp. 261–280.

Krashen, S. (1982), *Principles and practice in second language acquisition*, Pergamon Press, Oxford.

Krüger, A. & Hamilton, S. (1997), 'RECALL: Individual language tutoring through intelligent error diagnosis', *ReCALL* **9**(2), 51–58.

Kubota, M. (1994), 'The role of negative feedback on the acquisition of the English dative alternation by Japanese college students of EFL', *Institute for Research in Language Teaching Bulletin* **8**, 1–36. Available online from the Education Resources Information Center Database at http://www.eric.ed.gov.

Kubota, M. (1997), 'Instructional Effects of Positive and Negative Evidence on Prepositional/Phrasal Verbs', *Institute for Research in Language Teaching Bulletin* (11), 1–39. Available online from the Education Resources Information Center Database at http://www.eric.ed.gov.

Kumpf, L. (1984), Temporal systems and universality in interlanguage: A case study, *in* 'Universals of second language acquisition', Newbury House, pp. 132–143.

Kurup, M., Greer, J. E. & McCalla, G. I. (1992), The Fawlty Article Tutor, *in* C. Frasson, G. Gauthier & G. I. McCalla, eds, 'Intelligent Tutoring Systems', Springer Verlag, Berlin, pp. 84–91.

Labrie, G. & Singh, L. P. S. (1991), 'Parsing, Error Diagnostics and Instruction in a French Tutor', *CALICO Journal* **9**(1), 9–25.

Lafford, B. A. (1996), The development of tense/aspect relations in L2 Spanish narratives: Evidence to test competing theories. Paper presented at the Second Language Research Forum, October, Tucson, AZ.

Leeman, J. (2003), 'Recasts and Second Language Development: Beyond Negative Evidence', *Studies in Second Language Acquisition* **25**, 37–63.

197

Leiria, I. (1994), 'Aquisição do aspecto verbal por falantes nao-nativos de Portuguêse-Europeu: O exemplo dos pretéritos perfeito e imperferfeito [Acquisition of verbal aspect by nonnative speakers of European Portuguese: An example of perfect and imperfect past]', *Revista Internacional de Língua Portuguesa* **11**, 74–112.

Leiria, I. & Mendes, A. Q. (1995), Acquisition of tense and aspect in European Portuguese (L1 and L2), *in* I. S. Faria & M. J. Freitas, eds, 'Studies on the acquisition of Portuguese', Edições Colibri - Associação Portuguesa de Linguística, Lisbon, pp. 97–113.

Levy, M. (1997), *Computer-Assisted Language Learning: Context and Conceptualization*, Clarendon Paperbacks, Oxford, UK.

Levy, M. (2000), 'Scope, goals and methods in CALL research: Questions of coherence and autonomy', *ReCALL* **12**(2), 170–195.

L'haire, S. (2004), Vers un feedback plus intelligent: Les enseignements du projet Freetext, *in* 'Actes de la Journée d'étude de l'ATALA. TAL et Apprentissage des langues', Grenoble, pp. 1–12. Available at http://www.latl.unige.ch/freetext/en/publications.html.

Li, P. & Shirai, Y. (2000), *The acquisition of lexical and grammatical aspect*, Vol. 16 of *Studies on language acquisition*, Mouton de Gruyter, Berlin/New York.

Liceras, J. (1989), On some properties of the 'pro-drop' parameter: Looking for missing subjects in non-native Spanish, *in* S. Gass & J. Schachter, eds, 'Linguistic Perspectives on Second Language Acquisition', Cambridge University Press, Cambridge, pp. 109–133.

Liskin-Gasparro, J. (1997), Acquisition of tense and aspect in Spanish: Exploring learners' perceptions. Paper presented to the American Association of Applied Linguistics, Orlando, FL.

Liu, M., Moore, Z., Graham, L. & Lee, S. (2003), 'A Look at the Research on Computer–Based Technology Use in Second Language Learning: A Review of the Literature from 1990–2000', *Journal of Research on Technology in Education* **34**(3), 250–273.

Long, M. H. (1981), Input, interaction, and second language acquisition, *in* H. Winitz, ed., 'Native language and foreign language acquisition', Annals of the New York Academy of Sciences, 379, pp. 259–278.

Long, M. H. (1983), 'Linguistic and conversational adjustments to non-native speakers', *Studies in Second Language Acquisition* **5**, 177–194.

Long, M. H. (1988), Instructed interlanguage development, *in* L. Beebe, ed., 'Issues in Second Language Acquisition: Multiple Perspectives', Newbury House, Rowley, MA, pp. 115–141.

Long, M. H. (1991), Focus on form: A design feature in language teaching methodology, *in* K. de Bot, R. Ginsberg & C. Kramsch, eds, 'Foreign language research in cross-cultural perspective', John Benjamins, Amsterdam, pp. 39–52.

Long, M. H. (1996), The role of the linguistic environment in second language acquisition, *in* W. Ritchie & T. Bhatia, eds, 'Handbook of Second Language Acquisition', Academic Press, New York, pp. 413–468.

Long, M. H. & Doughty, C. J. (2003), SLA and Cognitive Science, *in* C. J. Doughty & M. H. Long, eds, 'The Handbook of Second Language Acquisition', Blackwell Publishing Ltd, Oxford, UK, pp. 866–870.

Long, M. H., Inagaki, S. & Ortega, L. (1998), 'The Role of Implicit Negative Feedback in SLA: Models and Recasts in Japanese and Spanish', *The Modern Language Journal* **82**(3), 357–371. Special Issue: The Role of Input and Interaction in Second Language Acquisition.

Loschky, L. (1994), 'Comprehensible input and second language acquisition: What is the relationship?', *Studies in Second Language Acquisition* **16**(3), 303–323.

Lyster, R. & Ranta, L. (1997), 'Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms', *Studies in Second Language Acquisition* **19**(1), 37–66.

Mackey, A. & Philp, J. (1998), 'Conversational Interaction and Second Language Development: Recasts, Responses, and Red Herrings?', *The Modern Language Journal* **82**(3), 338–356. Special Issue: The Role of Input and Interaction in Second Language Acquisition.

Malle, L. (1993), *Conversations avec Louis Malle/Philip French*, Denoël, Paris.

Mazurkewich, I. (1984), 'The Acquisition of the Dative Alternation by Second Language Learners and Linguistic Theory', *Language Learning* **34**(1), 91–109.

Meisel, J. M. (1987), Reference to Past Events and Actions in the Development of Natural Second Language Acquisition, *in* C. W. Pfaff, ed., 'First and Second Language Acquisition Processes', Newbury House Publishers Inc, Cambridge, pp. 206–224.

Moens, M. (1987), Tense, Aspect and Temporal Reference, PhD thesis, University of Edinburgh, Edinburgh, Scotland. Unpublished.

Moens, M. & Steedman, M. (1988), 'Temporal ontology and temporal reference', *Computational Linguistics* **14**, 15–28.

Montrul, S. & Slabakova, R. (2002), Acquiring morphosyntactic and semantic properties of aspectual tenses in L2 Spanish, *in* 'The acquisition of Spanish morpho–syntax: The L1/L2 connection', Kluwer, pp. 131–149.

Montrul, S. & Slabakova, R. (2003), 'Competence similarities between native and near–native speakers: An investigation of the preterite–imperfect contrast in Spanish', *Studies in Second Language Acquisition* **25**, 351–398.

Mourelatos, A. (1978), 'Events, Processes and States', *Linguistics and Philosophy* **2**, 415–434.

Murphy, M. & McTear, M. (1997), Learner Modelling for Intelligent CALL, *in* A. Jameson, C. Paris & C. Tasso, eds, 'User Modelling: Proceedings of the Sixth International Conference, UM97', Springer, Vienna/New York. Available for purchase at http://um.org.

Nagata, N. (1993), 'Intelligent Computer Feedback for Second Language Instruction', *The Modern Language Journal* **77**(3), 330–339.

Nagata, N. (1995), 'An Effective Application of Natural Language Processing in Second Language Instruction', *CALICO Journal* **13**(1), 47–67.

Nagata, N. (1996), 'Computer vs. Workbook Instruction in Second Language Acquisition', *CALICO Journal* **14**(1), 53–75.

Nagata, N. (1997a), 'An experimental comparison of deductive and inductive feedback generated by a simple parser', *System* **25**(4), 515–534.

Nagata, N. (1997b), 'The Effectiveness of Computer-Assisted Metalinguistic Instruction: A Case Study in Japanese', *Foreign Language Annals* **30**(2), 187–200.

Nagata, N. (1998), 'The Relative Effectiveness of Production and Comprehension Practice in Second Language Acquisition', *Computer Assisted Language Learning* **11**(2), 153–177.

Nagata, N. & Swisher, V. M. (1995), 'A Study of Consciousness-Raising by Computer: The Effect of Metalinguistic Feedback on Second Language Learning', *Foreign Language Annals* **28**(3), 337–347.

Nelson, K. E. (1987), Some observations from the perspective of the rare event cognitive comparison theory of language acquisition, *in* K. E. Nelson & A. vanKleeck, eds, 'Children's language', Vol. 6, Erlbaum, Mahwah, NJ, pp. 289–331.

Nobuyoshi, J. & Ellis, R. (1993), 'Focused communication tasks and second language acquisition', *ELT Journal* **47**(3), 203–210.

Norris, J. M. & Ortega, L. (2000), 'Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis', *Language Learning* **50**(3), 417–528.

Noyau, C. (1990), The development of means for temporality in the unguided acquisition of L2: Cross-linguistic perspectives, *in* H. W. Dechert, ed., 'Current trends in European second language acquisition research', Multilingual Matters, Clevedon, UK, pp. 143–170.

Noyau, C., Houdaïfa, E.-T., Vasseur, M.-T. & Véronique, D. (1995), The Acquisition of French, *in* 'The Acquisition of Temporality in a Second Language', John Benjamins, Amsterdam/Philadelphia, pp. 145–209.

Ortega, L. & Long, M. H. (1997), 'The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish', *Spanish Applied Linguistics* **1**, 65–86.

Piñón, C. (1995), An Ontology for Event Semantics, PhD thesis, Stanford University. Unpublished.

Pica, T. (1988), 'Interlanguage adjustments as an outcome of NS-NNS negotiated interaction', *Language Learning* **38**, 45–73.

Pica, T., Doughty, C. & Young, R. (1986), 'Making input comprehensible: Do interactional modifications help?', *ITL Review of Applied Linguistics* **72**, 1–25.

Pica, T., Young, R. & Doughty, C. (1987), 'The impact of interaction on comprehension', *TESOL Quarterly* **21**, 737–758.

Pienemann, M. (1998), *Language processing and second language development: Processability Theory*, Vol. 15 of *Studies in Bilingualism*, John Benjamins, Amsterdam/Philadelphia.

Pienemann, M. (2005), An Introduction to Processability Theory, *in* M. Pienemann, ed., 'Cross-Linguistic Aspects of Processability Theory', Vol. 30 of *Studies in Bilingualism*, John Benjamins, Amsterdam/Philadelphia, pp. 1–60.

Pienemann, M. & Johnston, M. (1987), Factors influencing the development of language proficiency, *in* D. Nunan, ed., 'Applying second language acquisition research', National Curriculum Resource Centre, AMEP, Adelaide, Australia, pp. 45–141.

Pienemann, M., Johnston, M. & Brindley, G. (1988), 'Constructing an acquisition-based procedure for second language assessment', *Studies in Second Language Acquisition* **10**, 217–243.

Pinker, S. (1984), *Language learnability and language development*, Harvard University Press, Cambridge, MA.

Pinker, S. (1989), *Learnability and cognition: The acquisition of argument structure*, MIT Press, Cambridge, MA.

Ramsay, V. (1990), Developmental stages in the acquisition of the perfective and the imperfective aspects by classroom L2 learners of Spanish, PhD thesis, University of Oregon, Eugene, OR. Unpublished.

Robison, R. E. (1990), 'The primacy of aspect: Aspectual marking in English interlanguage', *Studies in Second Language Acquisition* **12**, 315–330.

Robison, R. E. (1995), 'The Aspect Hypothesis Revisited: A Cross–Sectional Study of Tense and Aspect Marking in Interlanguage', *Applied Linguistics* **16**(3), 344–370.

Rohde, A. (1996), 'The aspect hypothesis and emergence of tense distinction in naturalistic L2 acquisition', *Linguistics* **34**, 1115–1138.

Salaberry, M. R. (2000), *The development of past tense morphology in L2 Spanish*, Vol. 22 of *Studies in bilingualism*, John Benjamins, Amsterdam/Philadelphia.

Salaberry, R. (1998), 'The development of aspectual distinctions in academic L2 French', *Canadian Modern Language Review* **54**, 508–542.

Salaberry, R. (1999), 'The development of past tense verbal morphology in classroom L2 Spanish', *Applied Linguistics* **20**, 151–178.

Salaberry, R. (2003), 'Tense aspect in verbal morphology', *Hispania* **86**, 559–573.

Salaberry, R. & Shirai, Y. (2002), L2 acquisition of tense-aspect morphology, *in* R. Salaberry & Y. Shirai, eds, 'The L2 Acquisition of Tense-Aspect Morphology', Vol. 27 of *Language Acquisition and Language Disorders*, John Benjamins, Amsterdam/Philadelphia, pp. 1–20.

Sanz, C. (1997), Experimental tasks in SLA research: Amount of production, modality, memory, and production processes, *in* A. T. Pérez-Leroux & W. R. Glass, eds, 'Contemporary Perspectives on the Acquisition of Spanish', Vol. Volume 2, Production, Processing, and Comprehension, Cascadilla Press, Somerville, MA, pp. 41–56.

Sanz, C. (2004), Computer Delivered Implicit Vs. Explicit Feedback in Processing Instruction, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 241–255.

Saxton, M. (1997), 'The Contrast Theory of negative input', *Journal of Child Language* **24**, 139–161.

Schachter, J. (1990), 'On the issue of completeness in second language acquisition', *Second Language Research* **6**, 93–124.

Schachter, J. (1991), 'Corrective feedback in historical perspective', *Second Language Research* **7**(2), 89–102.

Schachter, J. (1996), Maturation and the Issue of Universal Grammar in Second Language Acquisition, *in* W. Ritchie & T. Bhatia, eds, 'Handbook of Second Language Acquisition', Academic Press, San Diego, CA, pp. 159–189.

Schulze, M. (1999), 'From the developer to the learner: Describing grammar - learning grammar', *ReCALL* **11**(1), 117–124.

Schulze, M. & Hamel, M.-J. (2000), 'Towards authentic tasks and experiences: The example of parser–based CALL', *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée* **3**(1/2), 79–90.

Schumann, J. (1987), 'The expression of temporality in basilang speech', *Studies in Second Language Acquisition* **9**, 21–41.

Schwartz, B. (1993), 'On explicit and negative data effecting and affecting competence and linguistic behaviour', *Studies in Second Language Acquisition* .

Schwartz, B. D. (1986), 'The epistemological status of second language acquisition', *Second Language Research* **2**(2), 120–159.

Schwartz, B. D. (1987), The modular basis of second language acquisition, PhD thesis, University of Southern California. Unpublished.

Schwartz, B. & Sprouse, R. (1996), 'L2 cognitive states and the Full Transfer/Full Access Hypothesis', *Second Language Research* **12**, 40–72.

Schwind, C. B. (1990), 'An intelligent language tutoring system', *International Journal of Man-Machine Studies* **33**, 557–579.

Seliger, H. W. (1989), Deterioration and creativity in childhood bilingualism, *in* K. Hyltenstam & L. K. Obler, eds, 'Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss', Cambridge University Press, Cambridge, UK, pp. 173–184.

Seliger, H. W. (1991), Language attrition, reduced redundancy and creativity, *in* H. W. Seliger & R. M. Vago, eds, 'First language attrition', Cambridge University Press, Cambridge, UK, pp. 227–240.

Shirai, Y. (1991), Primacy of aspect in language acquisition: Simplified input and prototype, PhD thesis, University of California, Los Angeles. Unpublished.

Shirai, Y. (1995), Tense–aspect marking by L2 learners of Japanese, *in* D. MacLaughlin & S. McEwen, eds, 'Proceedings of 19th Annual Boston

University Conference in Language Development', Vol. 2, Cascadilla Press, Somerville, MA, pp. 575–586.

Shirai, Y. & Kurono, A. (1998), 'The acquisition of tense–aspect marking in Japanese as a second language', *Language Learning* **48**, 245–279.

Slabakova, R. (1999), 'The Parameter of Aspect in Second Language Acquisition', *Second Language Research* **15**(3), 283–317.

Slabakova, R. (2002), 'Recent research on the acquisition of aspect: An embarrassment of riches?', *Second Language Research* **18**(2), 172–188.

Slabakova, R. (2005), 'What is so difficult about telicity marking in L2 Russian?', *Bilingualism: Language and Cognition* **8**(1), 63–77.

Slabakova, R. & Montrul, S. (2000), Acquiring semantic properties of preterite and imperfect tenses in L2 Spanish, *in* S. Howell, S. Fish & T. Keith-Lucas, eds, 'Proceedings of the 24th Boston University Conference on Language Development', Cascadilla Press, Somerville, MA, pp. 534–545.

Slabakova, R. & Montrul, S. (2002), On viewpoint aspect interpretation and its L2 acquisition: A UG perspective, *in* R. Salaberry & Y. Shirai, eds, 'The L2 Acquisition of Tense-Aspect Morphology', Vol. 27 of *Language Acquisition and Language Disorders*, John Benjamins, Amsterdam/Philadelphia, pp. 363–395.

Slabakova, R. & Montrul, S. (2005), Acquisition at the Grammar-Discourse Interface: Aspectual Shifts in L2 Spanish, *in* J. Liceras, H. Zobl & H. Goodluck, eds, 'The Role of Features in Second Language Acquisition', Lawrence Erlbaum Associates, Mahwah, New Jersey.

Smith, C. S. (1991), *The Parameter of Aspect*, Kluwer Academic Publishers, Dordrecht/Boston.

Sten, H. (1952), *Les Temps du Verb Fini (Indicatif) en Français Moderne*, Munksgaard, Copenhagen.

Swain, M. & Carroll, S. (1987), The immersion observation study, *in* B. Harley, P. Allen, J. Cummins & M. Swain, eds, 'The development of bilingual proficiency final report: Classroom treatment', Vol. 2, Modern Language Centre, The Ontario Institute for Studies in Education, Toronto, pp. 190–342.

Taylor, B. (1977), 'Tense and Continuity', *Linguistics and Philosophy* **1**, 199–220.

Trévise, A. (1987), Toward an analysis of the (inter)language activity of referring to time in narratives, *in* C. W. Pfaff, ed., 'First and second language acquisition processes', Newbury House, Cambridge, MA, pp. 225–251.

Twaddell, W. F. (1965), *The English Verb Auxiliaries*, Brown University Press, Providence.

VanPatten, B. (1993), 'Grammar teaching for the acquisition rich classroom', *Foreign Language Annals* **26**, 435–450.

VanPatten, B. (1996), *Input Processing and Grammar Instruction in Second Language Acquisition*, Ablex Publishing Corporation, Norwood, New Jersey.

VanPatten, B. (2002a), 'Processing Instruction: An Update', *Language Learning* **52**(4), 755–803.

VanPatten, B. (2002b), 'Processing instruction, prior awareness, and the nature of second language acquisition: A (partial) response to Batstone', *Language Awareness* **11**, 240–258.

VanPatten, B. (2004a), Input Processing in Second Language Acquisition, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 5–31.

VanPatten, B. (2004b), Input and Output in Establishing Form–Meaning Connections, *in* B. VanPatten, J. Williams, S. Rott & M. Overstreet, eds, 'Form–Meaning Connections in Second Language Acquisition', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 29–47.

VanPatten, B. & Cadierno, T. (1993), 'Explicit instruction and input processing', *Studies in Second Language Acquisition* **15**(2), 225–243.

VanPatten, B. & Fernández, C. (2004), The Long–Term Effects of Processing Instruction, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 273–289.

VanPatten, B., Lee, J. F., Glass, W. R. & Binkowski, D. D. (1992), *Manual to Accompany ¿Sabías que...?*, McGraw Hill, New York.

VanPatten, B. & Oikkenon, S. (1996), 'The causative variables in processing instruction: Explanation vs. structured input activities', *Studies in Second Language Acquisition* **18**(4), 495–510.

VanPatten, B. & Sanz, C. (1995), From input to output: Processing instruction in communicative tasks, *in* F. R. Eckman, D. Highland, P. W. Lee, J. Mileham & R. R. Weber, eds, 'Second Language Acquisition Theory and Pedagogy', Lawrence Erlbaum Associates, Hillsdale, NY, pp. 169–186.

VanPatten, B. & Wong, W. (2004), Processing Instruction and the French Causative: Another Replication, *in* B. VanPatten, ed., 'Processing Instruction:

Theory, Research, and Commentary', Lawrence Erlbaum, Mahwah, NJ, pp. 97–118.

Varonis, E. & Gass, S. (1985), 'Non-native/non-native conversations: A model for negotiation of meaning', *Applied Linguistics* **6**(71–90).

Vendler, Z. (1967), *Linguistics in Philosophy*, Cornell University Press, Ithaca, NY.

Verkuyl, H. J. (1972), *On the Compositional Nature of the Aspects*, Vol. 15 of *Foundations of Language Supplementary Series*, Reidel, Dordrecht.

Verkuyl, H. J. (1989), 'Aspectual classes and aspectual composition', *Linguistics and Philosophy* **12**, 39–94.

Verkuyl, H. J. (1993), *A Theory of Aspectuality: The Interaction between Temporal and Attemporal Structure*, Cambridge University Press, Cambridge, UK.

Véronique, D. (1987), Reference to Past Events and Actions in Narratives in L2: Insights from North African Workers' French, *in* C. W. Pfaff, ed., 'First and second language acquisition processes', Newbury House, Cambridge, MA, pp. 252–272.

Vet, C. (1980), *Temps: Aspects et Adverbes de Temps en Français Contemporain*, Librairie Droz, Genève.

von Stutterheim, C. (1986), *Temporalität in der Zweitsprache: eine Untersuchung zum Erwerb des Deutschen durch türkische Gastarbeiter*, W. de Gruyter, Berlin/New York.

von Stutterheim, C. & Klein, W. (1987), A Concept-Oriented Approach to Second Language Studies, *in* C. W. Pfaff, ed., 'First and Second Language Acquisition Processes', Newbury House Publishers Inc, Cambridge, pp. 191–205.

Wang, Y. & Garigliano, R. (1992), An Intelligent Language Tutoring System for Handling Errors Caused by Transfer, *in* C. Frasson, G. Gauthier & G. I. McCalla, eds, 'Intelligent Tutoring Systems', Lecture Notes in Computer Science, Springer Verlag, pp. 395–404.

Weist, R. M., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E. & Konieczna, E. (1984), 'The defective tense hypothesis: On the emergence of tense and aspect in child Polish', *Journal of Child Language* **11**, 347–374.

White, L. (1985), 'The Pro-Drop Parameter in Adult Second Language Acquisition', *Language Learning* **35**, 47–62.

White, L. (1986), Implications of Parametric Variation for Adult Second Language Acquisition: An Investigation of the 'Pro-Drop' Parameter, *in* V. Cook,

ed., 'Experimental Approaches to Second Language Acquisition', Pergamon, Oxford, pp. 55–72.

White, L. (1987), 'Markedness and second language acquisition: The question of transfer', *Studies in Second Language Acquisition* **9**(3), 261–285.

White, L. (1991), 'The Verb-Movement Parameter in Second Language Acquisition', *Language Acquisition* **1**, 337–360.

White, L. (1992), 'Long and Short Verb Movement in Second Language Acquisition', *Canadian Journal of Linguistics* **27**, 273–286.

Wiberg, E. (1996), 'Reference to past events in bilingual Italian-Swedish children of school age', *Linguistics* **34**, 1087–1114.

Wolfram, W. (1985), 'Variability in tense marking: A case for the obvious', *Language Learning* **35**, 229–253.

Wolfram, W. (1989), Systematic variability in second-language tense marking, *in* M. R. Eisenstein, ed., 'The dynamic interlanguage: Empirical studies in second language variation', Plenum, New York, pp. 187–197.

Wong, W. (2004), Processing Instruction in French: The Roles of Explicit Information and Structured Input, *in* B. VanPatten, ed., 'Processing Instruction: Theory, Research and Commentary', Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 187–205.

Yang, J. C. & Akahori, K. (1997), Development of computer assisted language learning system for Japanese writing using natural language processing techniques: A study on passive voice, *in* 'Proceedings of the workshop "Intelligent Educational Systems on the World Wide Web", 8th World Conference of the AIED Society, Kobe, Japan, August 18–22, 1997'. Available at http://www.contrib.andrew.cmu.edu/~plb/AIED97_workshop/Yang/Yang.html.

Yoshii, R. & Milne, A. (1998), 'DaRT: A CALL System to Help Students Practice and Develop Reasoning Skills in Choosing English Articles', *CALICO Journal* **16**(2), 121–155.

Zock, M., Laroui, A. & Francopoulo, G. (1989), 'SWIM: A "Natural" Interface for the Scientifically Minded Language Learner', *Computers and the Humanities* **23**, 411–422.