Form-focused instruction and the acquisition of tense by Dutch-speaking learners of English: Experimental studies into the effects of input practice and output practice
Ureel, J.J.J.

Citation for published version (APA):
Form-focused instruction and the acquisition of tense by Dutch-speaking learners of English

Experimental studies into the effects of input practice and output practice

What is the difference between *I have done a lot of work this morning* and *I did a lot of work this morning*? Are sentences such as *I have seen him yesterday* and *Who has invented the telephone?* gramatically correct in present-day English? These are a few examples of the challenges that learners and teachers of English as a second language face when acquiring and teaching tense in English in instructional settings (e.g., grammar classes).

The concept of practice as a necessary component for learning a language is a given for many learners and teachers. However, it has received relatively little attention from a theory-based perspective. In the past, much of the instructional practice related to tense in English consisted of mainly output-focused drills, which required learners to produce grammatically correct forms. Although this approach has found its way into contemporary language teaching and has proven useful in its own ways, it has focused almost exclusively on output practice to the exclusion of other forms of practice. Contemporary materials designed for explicit instruction on tense in English have seen changes such as the addition of input-based practice, which requires learners to interact with input without producing any forms.

This dissertation addresses the concept of practice by investigating the effects of input practice and output practice on the acquisition of tense by intermediate Dutch-speaking learners of English. It does so by comparing an input processing hypothesis with skill acquisition theory in three computer-controlled learning experiments, involving 216 learners of English.
FORM-FOCUSED INSTRUCTION AND THE ACQUISITION OF TENSE BY DUTCH-SPEAKING LEARNERS OF ENGLISH

Experimental studies into the effects of input practice and output practice
FORM-FOCUSED INSTRUCTION AND THE ACQUISITION OF TENSE BY DUTCH-SPEAKING LEARNERS OF ENGLISH

Experimental studies into the effects of input practice and output practice

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D. C. van den Boom
ten overstaan van een door het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag, 14 januari 2011, te 14:00 uur

door

Jimmy John Jozef Ureel

geboren te Antwerpen, België
Promotiecommissie

Promotor: Prof. dr. J. H. Hulstijn
Co-promotor: Dr. L. Heyvaert

Overige leden: Prof. dr. R. M. DeKeyser
               Prof. dr. O. C. M. Fischer
               Prof. dr. M. Hannay
               Prof. dr. F. Kuiken

Faculteit der Geesteswetenschappen
To my mother, Elaine,
and my grandmother, JoJo

with love ... the *strong sort*
CONTENTS

Acknowledgements xi

Figures and tables xiii

Abbreviations and symbols xvii

Introduction 1

1 Tense and linguistics 9

1.1 Introduction 9
1.2 Defining tense 9
1.3 Conceptualizing tense 20
  1.3.1 Linguistic time in English 20
  1.3.2 Relative tense: The expression of tense relations within a domain 22
1.4 Locating bygone situations: The past and present perfect from a crosslinguistic perspective 23
  1.4.1 The past and present perfect in English 23
  1.4.2 The past and present perfect in Dutch 29
1.5 Conclusion 35

2 Tense and complexity 39

2.1 Introduction 39
2.2 Operationalizing temporal SLA verb-phrase complexity 40
2.3 Complexity: A multifaceted concept 42
  2.3.1 Linguistic complexity 42
    2.3.1.1 Absolute versus relative linguistic complexity 44
    2.3.1.2 Global versus local linguistic complexity 46
  2.3.2 Complexity in SLA research 47
    2.3.2.1 Complexity as a dependent or independent variable 47
    2.3.2.2 The absolute–relative linguistic complexity distinction applied in SLA research 50
  2.3.2.3 Complexity and the acquisition of L2 mappings 51
## 2.4 Relative, objective and psycholinguistic temporal SLA complexity and the English verb phrase

- **2.4.1 Introduction** 53
- **2.4.2 The complex L2 target features under investigation** 54
- **2.4.3 Form-related complexity** 58
- **2.4.4 Meaning-related and use-related complexity** 65
- **2.4.5 Mapping-related complexity** 66
- **2.4.6 Crosslinguistic influence and outcome measure complexity as complexity-inducing factors** 74

## 2.5 Conclusion

79

## 3 Tense and instructed SLA

- **3.1 Introduction** 81
- **3.2 Investigating the L2 acquisition of temporal expression** 82
  - **3.2.1 Early investigations into temporal expression** 82
  - **3.2.2 The form-oriented approach** 83
  - **3.2.3 The meaning-oriented approach** 85
  - **3.2.4 Comparing form-oriented and meaning-oriented approaches** 90
- **3.3 Conceptualizing form-focused instruction** 95
  - **3.3.1 A conceptual framework for operationalizing form-focused instruction research** 96
    - **3.3.1.1 Focus on form/focus on forms** 96
    - **3.3.1.2 Questions raised in L2 type-of-instruction research** 100
  - **3.3.2 The effects of practice-based instruction according to input processing and skill acquisition theory** 101
    - **3.3.2.1 Input processing and processing instruction** 101
    - **3.3.2.2 Skill acquisition theory** 108
- **3.4 Conclusion** 113

## 4 Study 1

- **4.1 Introduction** 115
- **4.2 Method** 117
  - **4.2.1 Design** 117
  - **4.2.2 Participants** 118
  - **4.2.3 Apparatus and materials** 122
    - **4.2.3.1 Apparatus** 122
    - **4.2.3.2 Materials** 122
  - **4.2.4 Procedure** 133
  - **4.2.5 Coding and scoring** 135
## CONTENTS

4.2.6 Statistical analyses 136
4.2.7 Validity and reliability 137
4.3 Quantitative results 137
4.4 Conclusions and discussion 145
4.5 Implications for quantitative data collection in Study 2 146
4.6 Qualitative results 147

5 Study 2 151
5.1 Introduction 151
5.2 Method 152
  5.2.1 Design 152
  5.2.2 Participants 153
  5.2.3 Apparatus and materials 156
    5.2.3.1 Apparatus 156
    5.2.3.2 Materials 156
  5.2.4 Procedure 158
  5.2.5 Coding and scoring 159
  5.2.6 Statistical analyses 159
  5.2.7 Validity and reliability 161
5.3 Results 160
5.4 Conclusions and discussion 168
5.5 Implications for quantitative data collection in Study 3 168

6 Study 3 169
6.1 Introduction 169
6.2 Method 171
  6.2.1 Design 171
  6.2.2 Participants 172
  6.2.3 Apparatus and materials 175
    6.2.3.1 Apparatus 175
    6.2.3.2 Materials 175
  6.2.4 Procedure 177
  6.2.5 Coding and scoring 177
  6.2.6 Statistical analyses 177
  6.2.7 Validity and reliability 178
6.3 Results 178
6.4 Conclusions and discussion 191
# Contents

7 General discussion 193

- 7.1 Introduction 193
- 7.2 Summary of the findings 193
- 7.3 Discussion of the findings 196
- 7.4 Research strengths and limitations 202
- 7.5 Implications for language pedagogy 206
- 7.6 Suggestions for further research 207
- 7.7 Conclusion 209

References 211

Appendix 227

Summary in Dutch 237

Summary in English 243

Curriculum vitae 249
ACKNOWLEDGEMENTS

I owe a debt of gratitude to a whole range of people, who have all supported me in their own ways in writing this doctoral dissertation. That is why I would like to take this opportunity to thank everyone who has been involved—directly or indirectly—in completing this dissertation. There are some people who deserve a special thank you for their support, which they provided with their expertise, time and words of encouragement.

First of all, I would like to thank my supervisor, Jan Hulstijn, who was willing to accept me as an external PhD candidate when I approached him with my interest in second language acquisition and the instruction of grammar. In his overly busy schedule, he consistently found time to provide me with astute and meticulous comments and he taught me some of the virtues of professional time management. He informed me from the start that he would read any texts that I submitted from his own experiences and education as an SLA scholar, and would not necessarily comment in detail on the quality of the English. However, he had an amazing—some would even say uncanny—knack of noticing dangling modifiers in the texts that I submitted for revision. To this day, I have no idea how and where Jan’s dangling-modifier detection skills were honed but I am grateful to him for pointing them out to me and for all the support that he gave me during the highs and lows of my research project.

I would also like to thank my co-supervisor, Liesbet Heyvaert, who—as a result of circumstances beyond anyone’s control—accepted this research project relatively late. She was extremely rigorous in reading the manuscript and providing judicious and detailed feedback wherever necessary. Her remarks about cohesion and coherence were sometimes very direct yet completely fair, insightful and constructive. Integrating her feedback was challenging at times but terribly rewarding. I cannot thank her enough for accepting this research project at such short notice and for her total commitment to her role as my co-supervisor.

Because of the experimental nature of the research, I had to look for a computer programmer who was able to ‘translate’ the research design and methodology for the three studies into a practical, user-friendly, computer-aided environment. This is where Koen De Groot turned out to be a godsend. As a computer programmer with a foreign language background, he was extremely flexible in accepting and making suggestions. The result was a wonderfully set up, transparent and extremely user-friendly environment, in which the participants felt at ease during the experimental sessions. I know that our collaboration was not all plain sailing and that I may have come across as somewhat of a nitpicker. However, Koen was extremely relaxed about making necessary changes and accommodating any ‘anal retentive’ wishes that I expressed.
ACKNOWLEDGEMENTS

A word of thanks must also be directed at my paranyphs, Iris Schrijver and Frédéric Wille. Both willingly accepted the invitations that I extended to them to act as my paranyphs. In addition, they also provided comments on the contents and the form of the manuscript, for which I am extremely grateful.

There are three other groups of people that have been instrumental—from a professional point of view—in making this dissertation what it is. I would like to extend a heartfelt thank you to all three: (1) the CASLA research group at the University of Amsterdam, (2) my colleagues at the Department of Translators and Interpreters at Artesis University College Antwerp (Frank Peeters and Anne Verhaert in particular, for the financial support that they made available, which enabled me to complete this dissertation) and (3) all the 200 plus students who participated in the three studies that were carried out for this dissertation.

Some colleagues, scholars and friends helped me with feedback on specific aspects of this dissertation or with more personal support throughout the years that I was working on this dissertation. I would like to thank the following people in particular: Kathleen Bardovi-Harlig, Annemie Bonneux†, Renaat Declerck, Benno King, Katrien Lievois, Anne Quataert and Aline Remael.

On a more personal note, I would like to thank my own family, my partner’s family and my friends for their encouragement during the years that I was working on and—according to some—‘stressing’ about my research. They may not have been able to contribute academically but they were always interested enough to ask me about the progress that I was making. A special, light-hearted comment for a friend at the Dutch Tax and Customs Administration: I hope that you of all people will thoroughly enjoy the fruits of my labour, or what you have jokingly referred to in Dutch as my werkstukje.

My partner Mark is the last person to be mentioned. He has been one of the staunchest supporters of my research to date. The last couple of years have been stressful for both him and myself, with research deadlines, which had to be met. However, throughout the stressful times, I was always able to turn to him for support, understanding and, most important, welcome distraction. I am fully aware that I may have been a handful while working on this dissertation but the unconditional support—in all possible ways imaginable—that he has given me has been admirable. As much as I love languages and grammar rules, they are simply powerful instruments and convenient nuts and bolts for communication. They themselves do not communicate with you when you want to express happiness, joy, elation, sadness, frustration and exasperation. Mark does, has done and continues to do so. That is why I think the world of him.

Jim Ureel
Amsterdam, 8 November 2010
# FIGURES AND TABLES

**Figure – Introduction**

1.1 A three-dimensional grammar framework 6

**Figures – Chapter 1**

1.1 Nomenclature of a possible 16-tense verb system in present-day English 15
1.2 Conjugation of the verb *talk* in a possible 16-tense verb system in present-day English 16
1.3 Linguistic conceptualization of the time line in English 21
1.4 The location of bygone situations on the time line in English by means of the past and the present perfect 24
1.5 The representation of the Dutch past using the descriptive elements *point of the event*, *point of reference* and *point of speech* 31
1.6 The representation of the Dutch present perfect using the descriptive elements *point of the event*, *point of reference* and *point of speech* 31

**Figures – Chapter 2**

2.1 Form, meaning and use of the English past when used to locate the bygone situation in the sentence *She asked me a question* 56
2.2 Form, meaning and use of the English present perfect when used to locate the bygone situation in the sentence *She has asked me a question* 57
2.3 The first principle and six subprinciples of VanPatten’s input processing model 69

**Figures – Chapter 3**

3.1 A form-oriented approach to the simple past in present-day English 84
3.2 A meaning-oriented approach to the expression of ‘bygone-ness’ in present-day English 86
3.3 A comparison of form-oriented and meaning-oriented studies of temporal expression in SLA research 93
3.4 The multicomponential nature of SLA 103
3.5 An adaptation of the multicomponential nature of SLA 103
3.6 Developmental stages in skill acquisition theory 110
### Figures – Chapter 4

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Design box for Study 1 (3 x 2 x 2 x 4 mixed design)</td>
<td>118</td>
</tr>
<tr>
<td>4.2</td>
<td>Grammaticality judgement (GJ): Example text and task type</td>
<td>125</td>
</tr>
<tr>
<td>4.3</td>
<td>Selected response (SR): Example text and task type</td>
<td>126</td>
</tr>
<tr>
<td>4.4</td>
<td>Constrained constructed response (CCR): Example text and task type</td>
<td>127</td>
</tr>
<tr>
<td>4.5</td>
<td>Translation (TR): Example text and task type</td>
<td>128</td>
</tr>
<tr>
<td>4.6</td>
<td>Example text for input-only group</td>
<td>130</td>
</tr>
<tr>
<td>4.7</td>
<td>Picture selection task with feedback for input-practice group</td>
<td>131</td>
</tr>
<tr>
<td>4.8</td>
<td>Example text and translation task for output-practice group</td>
<td>132</td>
</tr>
<tr>
<td>4.9</td>
<td>Levels of statistical analysis</td>
<td>138</td>
</tr>
<tr>
<td>4.10</td>
<td>Performance on receptive pretest and posttest by treatment group</td>
<td>140</td>
</tr>
<tr>
<td>4.11</td>
<td>Performance on productive pretest and posttest by treatment group</td>
<td>140</td>
</tr>
<tr>
<td>4.12</td>
<td>Overall mean performance on receptive and productive pretest and posttest</td>
<td>141</td>
</tr>
<tr>
<td>4.13</td>
<td>Pretest and posttest performance on all four outcome measures by treatment group</td>
<td>144</td>
</tr>
<tr>
<td>4.14</td>
<td>Pretest and posttest mean performance by treatment group</td>
<td>145</td>
</tr>
</tbody>
</table>

### Figures – Chapter 5

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Design box for Study 2 (2 x 2 x 2 x 4 mixed design)</td>
<td>153</td>
</tr>
<tr>
<td>5.2</td>
<td>Example text and question for input-practice group</td>
<td>158</td>
</tr>
<tr>
<td>5.3</td>
<td>Performance on receptive pretest and posttest by treatment group</td>
<td>162</td>
</tr>
<tr>
<td>5.4</td>
<td>Performance on productive pretest and posttest by treatment group</td>
<td>162</td>
</tr>
<tr>
<td>5.5</td>
<td>Overall mean performance on receptive and productive pretest and posttest</td>
<td>163</td>
</tr>
<tr>
<td>5.6</td>
<td>Interaction between overall mean performance on test items and time of testing</td>
<td>163</td>
</tr>
<tr>
<td>5.7</td>
<td>Pretest and posttest performance on all four outcome measures by treatment group</td>
<td>166</td>
</tr>
<tr>
<td>5.8</td>
<td>Pretest and posttest mean performance by treatment group</td>
<td>167</td>
</tr>
<tr>
<td>5.9</td>
<td>Interaction between overall mean performances on outcome measures and time of testing</td>
<td>167</td>
</tr>
</tbody>
</table>

### Figures – Chapter 6

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Design box for Study 3 (3 x 3 x 2 x 4 mixed design)</td>
<td>172</td>
</tr>
<tr>
<td>6.2</td>
<td>Performance on receptive pretest, immediate posttest and delayed posttest by treatment group</td>
<td>182</td>
</tr>
</tbody>
</table>
FIGURES AND TABLES

6.3 Performance on productive pretest, immediate posttest and delayed posttest by treatment group 183
6.4 Interaction between time of testing and treatment 184
6.5 Interaction between time of testing and modality 184
6.6 Pretest, immediate posttest and delayed posttest performance on all four outcome measures by treatment group 189
6.7 Performance by time of testing and outcome measure 190

Tables – Chapter 1

1.1 Distribution of analytic and synthetic future verb forms in Germanic and Romance languages 18
1.2 Zone-specific categorization of time-specifying adverbials 29

Tables – Chapter 3

3.1 Types of form-focused instruction 98
3.2 Main principles of input processing 105

Tables – Chapter 4

4.1 Language groups (bachelor in applied linguistics with a specialization in translation) 120
4.2 Participants’ language combinations (Study 1) 121
4.3 Performance on receptive and productive pretest and posttest items by treatment group 139
4.4 Performance on pretest and posttest items by outcome measure and by treatment group 142-143

Tables – Chapter 5

5.1 Participants’ language combinations (Study 2) 155
5.2 Performance on receptive and productive pretest and posttest items by treatment group 161
5.3 Performance on pretest and posttest items by outcome measure and by treatment group 164-165

Tables – Chapter 6

6.1 Participants’ language combinations (Study 3) 174
6.2 Performance on receptive and productive pretest and posttest items by treatment group 179-180
6.3 Performance on pretest and posttest items by outcome measure and by treatment group 185-187

Table – Chapter 7

7.1 Research questions addressed across Studies 1, 2 and 3 194
ABBREVIATIONS AND SYMBOLS

ACT adaptive control of thought
AmE American English
Anéla Association néerlandaise de linguistique appliquée
ANOVA analysis of variance
BrE British English
CALL computer-assisted language learning
CCR constrained constructed response (outcome measure)
CLI crosslinguistic influence
E point of the event (tense)
EI explicit information
EFL English as a foreign language
ESL English as a second language
FFI form-focused instruction
FMU (mapping) form–meaning–use (mapping)
GJ grammaticality judgement (outcome measure)
IP input processing
L1 first language
L2 second language
MOI meaning-based output instruction
vvt voltooid tegenwoordige tijd
PI processing instruction
R point of reference (tense)
RM ANOVA repeated-measures analysis of variance
S point of speech (tense)
SIA structured input activity
SR selected response (outcome measure)
SLA second language acquisition
t0 temporal zero-point
TAM tense-aspect-mood
TAP think-aloud protocol
TAP transfer-appropriate processing
TMA tense-mood-aspect
TR translation (outcome measure)
\( vtt \) voltooid tegenwoordige tijd
\( \alpha \) Cronbach’s alpha coefficient
The instruction of tense

Je suis, tu es, il est, nous sommes, vous êtes, ils sont. I can still remember the first time that I came across this short list of—what seemed to me at the time—challenging verb forms. The year was 1986 and I was taking my first compulsory French class at the age of ten. Back then, I did not think much of the forms. I simply remember rushing home and memorizing the list of forms for next week’s class. A couple of weeks later, once again during my French class, I came across another list of verb forms, which—surprisingly enough—did not bear much resemblance to the first: j’ai, tu as, il a, nous avons, vous avez, ils ont. Unconsciously, I had been trying to memorize the forms of the indicatif présent, the French present indicative, of the verbs être (be) and avoir (have). During the consecutive eight years of French classes, I came across many more lists of verb forms and—horror of horrors—a whole plethora of regular and irregular verb forms with possible meanings and uses which had to be mastered. It took me several years to figure out that what my teachers of French were expecting me to learn were actually tense forms (verb forms that generally express temporal meanings), mood forms (verb forms that generally express concepts such as certainty and uncertainty, obligation and possibility) and combinations of the two.

Teaching approaches may vary but at one point in our scholastic or academic careers in the Western world, most of us are required to take language classes and to master—in some shape or form—tense forms in the languages being learnt. English language classes are no exception in this respect and most English courses focus on tense at one point during the progression through a syllabus. The linguist Geoffrey Leech (2004) starts the third edition of his book Meaning and the English Verb by stating the following about tense:

Every language has its peculiar problems of meaning for the foreign language learner. Many people would agree that in the English language, some of the most troublesome yet fascinating problems are concentrated in the area of the finite verb phrase, including, in particular, tense [emphasis added], aspect, mood and modality. (p. 1)

In 2001, fifteen years after my first encounter with tense during my compulsory French classes, I found myself once again confronted with tense but this time as a beginning lecturer of English grammar in higher education. My target audience consisted of mainly Dutch-speaking students of English, most of whom were aspiring to become translators and interpreters. As the years progressed, I started to notice some recurring grammatical problems that my students, who were intermediate to advanced learners of English, tried to wrap their heads around in class. The grammatical category of tense was one of them. Generally, when I ask students of English what they think about the difficulty of English
INTRODUCTION

grammar, many reply that English is not that difficult but that the use of the tenses—whatever their definitions of tenses may be—in English can be problematic at times. It is this common reply which ignited my interest in the complexities of the English tense system and in the challenges that both learners and teachers of English as a second language (henceforth ESL learners and ESL teachers) face when dealing with temporal verbal forms in contexts in which explicit grammar instruction is the norm.

My two points of departure for this doctoral dissertation were the following: (1) the problematic nature of tense, which (Dutch-speaking) ESL learners inevitably have to grapple with and (2) the need—as an ESL teacher—to help learners overcome intractable temporal problems by providing them with help in an environment in which explicit grammar instruction is the norm. A traditional teaching approach to grammar—and to a whole range of other language features for that matter—has been the so-called PPP approach, where the three p’s stand for presentation, practice and production (Thornbury, 2003). In the past, much of the practice related to tense (in English) in the PPP approach consisted of mechanical drills, with ESL learners being required to fill in the right temporal form of infinitives provided in situational contexts. Although this approach to practice has found its way into more contemporary language teaching materials (e.g., Aarts & Wekker, 1993; Foley & Hall, 2003; Hoffmann & Hoffmann, 2001) and has proven useful in its own ways, it focuses almost exclusively on the production of output to the exclusion of any other forms of practice. Luckily, we have seen promising changes in materials designed for explicit instruction on tense in L2 English, with the addition of both situational and communicative materials based on, for example, various forms of input-based practice (e.g., grammaticality judgement, picture selection, selecting correct tense forms).

With these added materials come the issues of (1) how to incorporate such practice-based materials into settings in which explicit grammar instruction is the norm and (2) how to select materials based on the treatment effects that they may bring about. These questions encapsulate the essence of the research in this doctoral dissertation. With all the references to practice, one could think that the focus of this dissertation is purely on experimental research and that not much theory is involved in the incorporation of practice-based materials. However, a combination of both theoretical research and experimental research is required to be able to draw conclusions with respect to any acquisitionally informed pedagogy. Consequently, this dissertation will focus on both types of research.

This doctoral dissertation is the culmination of several years of research, which consisted of focusing on the instruction of one specific tense-related issue that appears to be highly problematic for many ESL learners of varying proficiency (i.e., beginner, intermediate, advanced): the distinction between the past and the present perfect when used to locate situations which have taken
place in the past and which are also referred to as bygone situations.1 Distinguishing between these two tenses when locating bygone situations is problematic and is generally an item of discussion in most descriptive and pedagogical grammars of English (e.g., Aarts & Wekker, 1993; Carter & McCarthy, 2006; Declerck, 1991, 2003, 2006; Foley & Hall, 2003; Koning & van der Voort, 1997; Leech, 2003). In addition, it is an issue which is also dealt with in most classes dealing with the English tense system. My research interest, however, was on how this problematic distinction may be dealt with from an instructional point of view through the incorporation of practice as a fundamental component. By focusing on the practice component, I wanted to shed light on the pedagogical implications of two forms of practice—input practice and output practice—on the acquisition of the problematic past/present perfect distinction.

Preview of the chapters
In Chapter 1, *Tense and linguistics*, the grammatical category of tense is placed in a linguistic context. By focusing on recurring linguistic issues in discussions of tense, I have tried to give the targeted past/present perfect distinction a linguistic dimension which takes into account the challenges related to discussing tense in general and the past/present perfect distinction in particular.

Chapter 2, *Tense and complexity*, builds a bridge between the purely linguistic aspects of tense complexity and the challenges that ESL learners—including Dutch-speaking ESL learners—face when trying to acquire tense in English. The chapter sheds light on the mainly qualitative features of tense complexity in an SLA context, which includes—but is not limited to—the discussion of linguistic issues of complexity. In addition to linguistic complexity, issues such as mapping complexity, SLA complexity and other complexity-inducing factors are also discussed.

In Chapter 3, *Tense and instructed SLA*, the focus is placed on tense in a context of instructed SLA. The study of temporality in SLA research came about incidentally but has developed into a well-structured, systematic and methodologically developed area of research. By first highlighting relevant features of studies into L2 temporality, I introduce to the reader the approaches that have been used to study L2 temporality. Subsequently, the focus is placed on instructed SLA and on two approaches that feature practice-based instructional setups as integral parts of the SLA process: (1) input processing and (2) skill acquisition theory. By describing the features of both approaches and by discussing the role of practice in both approaches, I give the reader the necessary background to be able to evaluate the experimental research carried out in later chapters.

---

1 For a more nuanced definition of the term *bygone*, see Chapter 1, Section 1.4.1.
Chapter 4, Study 1, introduces the first of three studies which are part of the empirical research carried out. By comparing the performances of three treatment groups (input only, input practice, output practice) during the pretest and posttest sessions of the study, the research questions which addressed any differential effects resulting from instructional treatment were investigated. Quantitative data were used to carry out analyses and based upon the results of those analyses, the conclusions drawn with respect to the effects of treatments were that there were no differential effects between the treatments. Consequently, four changes to the research design and methodology were formulated with a view to carrying out an improved second study.

Chapter 5, Study 2, covers the second study that was carried out one year later. The participants were new but the research questions were essentially the same as the ones formulated in Study 1. Having implemented the four changes that had been formulated when reporting on Study 1, I once again carried out quantitative analyses but this time with only two treatment groups (input practice, output practice) and with different forms of practice. When comparisons of the pretest and posttest data were carried out, the analyses showed a statistically significant improvement in the participants’ performances. However, once again no differential effects between treatments were found.

Chapter 6, Study 3, reports on the third and final study that was carried out. In effect, Study 3 was a replication of Study 2. The only change that was implemented was the addition of a control group to ascertain whether the instruction that was provided to the participants actually brought about significant changes in performance. This brought the number of treatment groups back to three (input practice, output practice, no treatment). The quantitative analyses showed significant increases in test performances for the input-practice and output-practice groups but not for the no-treatment group. However, a comparison of the data from the input-practice and output-practice groups revealed no significant differences between both groups. In other words, no differential effects between the two treatment groups were confirmed.

Chapter 7, General discussion, brings the theoretical and experimental chapters together by providing both a summary and a more detailed discussion of the findings from all three studies that were carried out. In addition, the strengths and limitations of the experimental research are highlighted. This is followed by a discussion of the three studies with respect to the implications of the results for language pedagogy. By way of conclusion, suggestions for further research are provided before a final conclusion is drawn.

Recurring themes and terminological issues

Form–meaning–use mapping. Although the idea of form-meaning mappings is an extremely common one in linguistics, it may come as a surprise to some that the traditional binary concept of form-meaning mapping has given way to a more extended ternary description in the title including not only form (mor-
INTRODUCTION

phosyntax) and meaning (semantics) but also use (pragmatics) (see Figure i.1). The concept of mapping is not new and has been in use—albeit without consistent and explicit references to the term mapping—in linguistics in general and in the field of second language acquisition in particular for quite some time. Over the years, the concept may have undergone some refinements but the approach of visualizing both the various dimensions of grammar and the interaction between those dimensions has proven extremely useful in the fields of language study and second language acquisition. One of the first linguists to introduce and actively promote a mapping approach to language was the Swiss linguist Ferdinand de Saussure. He himself took this idea from the field of semiotics and applied it to linguistics and, in so doing, was closely connected to the inception of the concept of linguistic signs (Finch, 2000). Saussure promoted the idea that a linguistic sign consisted of two elements: (1) a signifier (i.e., a pronunciation form) and (2) a signified (i.e., a meaning/sense) (de Saussure, 1983). The parallelism to the current concept of mapping in the fields of linguistics and second language acquisition is obvious and it is this concept of mapping which has been extensively used to represent language and the learning/teaching of language.

The concept of mapping is a basic concept in the field of (contemporary) SLA research (Saville-Troike, 2006). In the introductory chapter in the book Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology, editor Robert DeKeyser writes that “clearly, form–meaning connections are the essence of language” (2007a, p. 10). Consequently, one could assume that one of the tasks of SLA researchers is to investigate how these form–meaning connections are learnt in an L2. The idea and/or importance of mappings is one which is shared by (SLA) scholars and which is often—directly or indirectly—referred to in the (SLA) literature (e.g., Bardovi-Harlig, 2000; Bardovi-Harlig & Comajoan, 2008; Bates & MacWhinney, 1989; Brown, 2000; DeKeyser, 2005, 2007c; Ellis, 1997; Finch, 2000; Harmer, 2007; Housen, 2002; Larsen-Freeman, 2003; Salaberry & Shirai, 2002; Saville-Troike, 2006; Skehan, 1998; Thornbury, 1999, 2006; VanPatten, 2002a; VanPatten, Williams, Rott, & Overstreet, 2004). Thus, the ubiquity of and general references to (linguistic) mappings and the overall usefulness of a mapping approach were the two reasons why a mapping approach to grammar was chosen for the research presented here.

2 The concepts of form–meaning mapping and form–function mapping are often used interchangeably in the literature. However, strictly speaking, it could be argued that function entails more than simply meaning. In the case of grammar, knowing what a grammatical feature means (semantics) could be considered one aspect of the acquisition process whereas knowing how to use that feature (pragmatics) could pose an altogether different challenge for L2 learners. A ternary (or tripartite) approach to grammar (form–meaning–use) helps in highlighting possible distinctions between semantics and pragmatics. Such an approach may be used—beyond the discussion of tense—as a general approach to grammar and to grammar instruction in SLA contexts.
Another reason for choosing such an approach is that it can be viewed as approach-neutral/theory-neutral. The adjectives approach-neutral and theory-neutral refer to the idea that the tripartite approach to grammar can be used in a variety of (SLA) approaches and theories (ranging, for example, from approaches and theories with no focus on grammar to approaches and theories with a heavy focus on grammar) without having to worry about any theory-driven problems or inconsistencies. Whereas certain SLA approaches and theories clearly focus on grammar (often even on one specific aspect of grammar (i.e., form, meaning or use) and this to the exclusion of the other dimensions), the idea behind opting for a three-dimensional grammar framework is that such exclusion need not necessarily take place (Celce-Murcia & Larsen-Freeman, 1999). Proponents of a three-dimensional grammar framework (e.g., Celce-Murcia & Larsen-Freeman, 1999; DeCarrico & Larsen-Freeman, 2002; Larsen-Freeman, 2001, 2003) seek to promote the idea that all three dimensions of grammar should be embraced and that the study of grammar should not be reduced to the mere study of one dimension but should be undertaken studying all three dimensions. However, this need not be a requirement for using a ternary approach to grammar. A fundamental distinction should be made here between an approach to studying the object of study (i.e., grammar) and an approach to studying the method of how that object is studied. If it is assumed that the object (i.e., grammar) consists of three dimensions (form, meaning and use), there is no absolute need to embrace those three dimensions when undertaking the study, teaching and/or
learning of that object. For the research carried out in this doctoral dissertation, the three dimensions were important but the main focus was on meaning and use, which is explained in more detail in Chapters 1, 2 and 3 and put into practice in the three studies reported on in Chapters 4, 5 and 6. One deciding factor in the discussion of focus on one or several dimensions of grammar could be the ultimate goal of studying, teaching and/or learning the object of grammar. This then brings us back to the importance of the participants’ educational environment and the goal of the participants’ educational programme (see Chapters 4, 5 and 6). Accuracy, meaningfulness and appropriateness are three factors which influence translators during their educational development and their future careers. Consequently, this trichotomy fits in perfectly with a ternary mapping approach and grammar framework, especially if we take into consideration the educational context in which the participants of the studies found themselves, that is, a bachelor programme in applied linguistics with a specialization in translation.

**English as a second language/English as a foreign language.** In the field of second language acquisition, the concept of a second language often contrasts with the concept of a foreign language, with second-language status being awarded to a language which is learnt in a community where it is used as the usual language for communication (e.g., learning English in England) and foreign-language status being awarded to a language which is being learnt in a community where it is not the usual language for communication (e.g., learning English in Belgium or the Netherlands). With respect to English, this distinction has resulted in the concepts English as a second language (ESL) and English as a foreign language (EFL). However, the distinction between the two concepts is not systematically adhered to by many researchers. Consequently, the term second language is often used as a cover term for both a second language and a foreign language, regardless of the institutional and/or social role played by the language being learnt. In this dissertation, the term second language will be used as a cover term unless the aforementioned distinction is absolutely essential. This is in line with the belief that the distinction between ESL and EFL is regarded by many as an oversimplification (Nunan, 1999).

**Language acquisition/language learning.** Although many SLA scholars adhere to the language acquisition/language learning distinction (e.g., Krashen, 1981, 1987), this is not a distinction that I will adhere to in this doctoral dissertation. The distinction reflects the idea that language acquisition consists of an unconscious process by which a linguistic system (i.e., language) is internalized and language learning consists of a conscious process by which a linguistic system is internalized. In this doctoral dissertation, language acquisition and language learning will be used interchangeably to refer to the general process by which a linguistic system is internalized. The assumption in this dissertation is
that since all learning is—to some extent—cognitively controlled, the acquisition/learning distinction is not a distinction of kind, but rather a distinction of degree. No differentiated terminology will be used to refer to this distinction of degree.
CHAPTER 1
TENSE AND LINGUISTICS

Nothing is improbable until it moves into the past tense.
(George Ade)

1.1 Introduction
Many instructed foreign language learners who are asked to elaborate on the concept of tense appear to have some idea of what tense is. They are often able to cite paradigms of textbook examples of verbs in specific tenses in the languages that they are learning. However, providing a nuanced definition of the grammatical category of tense seems to be a different challenge altogether. The challenges involved in defining tense are many and they are reflected in the linguistic literature on the topic. The overall aim of this chapter is to contextualize the grammatical category of tense from a mainly linguistic point of view. I will first highlight some of the most common challenges which are generally featured in discussions of the grammatical category of tense. After this, I will discuss the meanings and uses of the past and the present perfect to locate bygone situations in English and in Dutch.

The first section of this chapter (Section 1.2) will look at some existing definitions of tense and will use these definitions to highlight and discuss the recurring items of agreement and disagreement that may be found in publications on tense. The second section (Section 1.3) will address the issue of how tense may be conceptualized (in English) by using an existing descriptive theory of tense in English. The third and final section (Section 1.4) will use the insights from the first two sections and will investigate the similarities and differences with respect to the past and the present perfect in English and Dutch.

1.2 Defining tense
The grammatical category of tense can be defined as a broad grammatical category in that it may be instantiated in language in a variety of ways and interacts with other grammatical categories (e.g., aspect, mood) and with a number of lexical and/or pragmatic means which are generally used to express temporality in language (e.g., adverbials, connectives, scaffolding). A closer look at a sele-
tion of definitions of the grammatical category of tense found in both the gen-
eral, introductory linguistic literature—applied and non-applied—and the more
specialized tense-related literature may help identify some of the most common
items of both agreement and disagreement among scholars.

Consider the following definitions of tense (quoted with original highlight-
ing):

**tense** *(adj./n.)* (1) *(tns, TNS)* A CATEGORY used in the GRAMMATICAL description
of VERBS (along with ASPECT and MOOD), referring primarily to the way the gram-
mar marks the time at which the action denoted by the verb took place. Traditionally, a
distinction is made between past, present and future tenses, often with further divi-
sions (perfect, pluperfect, etc.). In LINGUISTICS, the relationship between tense and
time has been the subject of much study, and it is now plain that there is no easily
stateable relationship between the two. Tense FORMS (i.e. variations in the MORPHO-
LOGICAL form of the verb) can be used to signal MEANINGS other than temporal
ones. In English, for example, the past-tense form *(e.g., I knew)* may signal a tentative
meaning, and not past time, in some contexts *(e.g., I wish I knew - that is, ‘know now’).*
Nor is there a simple one-to-one relationship between tense forms and time: the pre-
sent tense in English may help to refer to future or past time, depending on CONTEXT
*(e.g., I am going home tomorrow, Last week I am walking down this street . . .)* (see HISTORIC
PRESENT). Furthermore, if tenses are defined as forms of the verbs, it becomes a mat-
ter of debate whether a language like English has a future tense at all: constructions
such as *I will/shall go*, according to many, are best analysed as involving MODAL AUX-
ILIARY verbs, displaying a different grammatical FUNCTION *(e.g., the expression of
intention or obligation, which may often involve futurity).* English illustrates several
such problems, as do other languages, where tense forms, if they exist, regularly display
analytic difficulties, because of overlaps between tense and other verbal functions, such
as aspect and mood. Alternative terminology *(e.g., ‘past’ v. ‘non-past’, ‘future’ v. ‘non-
future’, ‘now’ v. ‘remote’)* will often be needed. *(Crystal, 2003, pp. 459–460)*

**Tense**: (a) the phenomenon that a language has a special system of verb forms to lo-
cate *(the actualizations of situations)* in time; (b) the correlation of a particular gram-
matical form with a particular tense meaning *(e.g., the ‘past tense’). In more detail:
tense is a linguistic concept *(as opposed to time)* denoting the form taken by the verb
to locate *(the actualization of)* the situation referred to in time, i.e. to express the tem-
poral relation between the time of the situation in question and an orientation time
which may be either the temporal zero-point or another orientation time that is tempo-
rally related *(directly or indirectly)* to the temporal zero-point. English has several
tenses, such as the present tense, the past tense, etc., to which correspond different
verb forms, which are called the tense forms of the verb. *(Declerck, 2006, p. 820)*

**Tense**. A category used in the description of VERBS which refers to the location of an
action in time *(as distinct from ASPECT, which is concerned with its duration).* Many
linguists make a distinction between form and function in analysing tense. By this
analysis there are only two tenses in English – past and present. In other words, there
are only two ways in which tense is grammaticalised. If we wish to indicate the present
tense we use the base form of the verb, whilst the past involved adding ‘ed’ *(in the case
of regular verbs).* There is no separate inflection for the future, however. To indicate
that, we use either the present tense, *I am going home tomorrow,* or the modal verbs *shall*
and will in conjunction with a main verb, *I shall/will go home tomorrow.* The distinction be-
tween form and function is a useful one in dealing with the complicated relationship
between tense and time. Tensed forms (i.e. variations in the structure, or morphology, of the verb) can be used to signal a wide variety of meanings other than temporal ones. We can, for example, use the past tense to indicate some kind of hypothetical meaning, as in, *I wish I had your money* (i.e. have now). . . . What is clear from studies of tense and time is that there is no easily stateable relationship between them. We use the two formal grammatical distinctions, past and present, to perform a number of functions, some of which are purely temporal, and some which are modal, or interpersonal in some way. (Finch, 2000, pp. 118–119)

**Tense**

The different tense forms of **verbs** are forms used primarily to express the time at which an event occurred, or at which some state of affairs held. It is conventional, in everyday talk about language, to assume that there are just three basic tenses, namely **past**, **present** and **future**. But when one looks at the different ways in which languages express distinctions related to time, one generally finds a more complicated picture, and often a picture that does not reflect the simple past/present/future division.

Traditional grammarians and modern linguists have approached this complicated area of languages with slightly different terminological conventions. What many traditional grammarians label as various kinds of ‘tense’, modern linguists split into two different ideas, namely: tense, which is strictly to do with **WHEN** something happened or was the case; aspect, which is concerned with factors such as the **DURATION** or **COMPLETENESS** of events and states of affairs. For English, this difference of terminology comes out mainly in relation to the **perfect** and the **progressive**, which many traditional grammarians would treat as part of the system of tense, but modern linguists treat as belonging to the system of aspect. In this dictionary, we keep to the traditional grammarian’s terminology. But ‘tense’ is an area in which the traditional terminology is indeed quite crude. The more modern distinction between tense and aspect is a valuable refinement, and advanced detailed work on languages must make this distinction. (Hurford, 1994, p. 239)

tense 1. A **grammatical category** which correlates fairly directly with time. Tense is usually, though not invariably, marked in the verb in those languages that have it. Some languages have only two tenses (usually past versus non-past, sometimes future versus non-future); some have three (usually past, present, future); and some have more (for example, they may distinguish recent past from remote past). The largest number of tenses so far reported is eleven, in the African language Bamileke-Dschang. English has only two tenses: past versus non-past (“present”), as in *loved/loves, went/go, would/will*. **(Thornbury, 2006, p. 226)**
Note: traditional grammarians often use the term ‘tense’ in a very loose way that includes also aspect and mood, but this usage is objectionable. (Trask, 1997, p. 218)

Note that some of the definitions above focus on defining the grammatical category of tense from a language-specific perspective—more specifically—an English-specific perspective. However, it should be borne in mind that choosing to focus on one specific language—or on a specific group of languages—when defining grammatical concepts may lead to definitional bias depending on the linguistic features of the language(s) being focused on. The following discussion of the items of agreement and disagreement will focus on English but will also, at times, highlight similarities and differences with other languages wherever relevant.

The first item of agreement that can be found in most of the definitions above is the general consensus that tense may be defined as a grammatical category of the verb. This may not be referred to in all of the definitions explicitly and by means of the linguistic term grammatical category but those definitions that make no explicit reference to the wording grammatical category do use ample examples of tense which show that tense is generally a grammatical category of the verb and that tense is expressed in the form of verbal paradigms. As a result, a large part of the discussion in this doctoral dissertation will focus on form-related, meaning-related and use-related features of temporality in the verb phrase. However, even though tense may be regarded as the “grammaticalised location in time” (Comrie, 1985, p. 9), it is one of several means which can be used to express temporality in language. In extended discourse, one generally finds the presence of several means of expressing temporality. This, in turn, usually results in nuanced forms of interplay between these means. A common example of such interplay is the interaction between the grammatical category of tense and, for example, (temporal) adverbials. The common presence of interplay explains why the other means of expressing temporality—especially the lexical means (e.g., adverbials)—will also be featured and discussed in this dissertation. It also explains why aspects of interplay will be highlighted wherever these are considered relevant for the development of ideas.

The second item of agreement is the reference to the relationship between tense and time. Most of the definitions above explicitly refer to a relationship between tense and time. At the same time, however, most of the definitions also refer to the fact that this relationship is not “easily stateable” (Crystal, 2003, p. 459) (i.e., is not straightforward) and list examples of the relationship.

---

5 In the majority of linguistic publications dealing with tense, the grammatical category of tense is defined as a grammatical category of the verb (e.g., Comrie, 1985; Leech, 2004; Lewis, 1986; Palmer, 1987). A probable reason for this method of approach is that tense is generally grammaticalized as a feature of the verb phrase in most languages that have tense (Comrie, 1985). However, even though tense is mainly approached as a grammatical category of the verb, it is not necessarily a feature that is exclusive to the verb phrase. Some languages also assign tense to, for example, adjectival phrases, adverbial phrases and noun phrases (Comrie, 1985).
One commonly discussed aspect of this relationship is that there is not always a one-to-one correspondence between tense and time. In other words, the relationship between tense and time is a bidirectionally fluid relationship. In practice, this means that one tense may express several temporal meanings and one temporal meaning may be expressed by several tenses. By way of example, let us have a closer look at the present tense in English and at the times that this tense may refer to. The prototypical or primary meanings of the present tense are meanings that refer to situations that hold at the present moment (e.g., *I am here now*, *Beckham kicks the ball to Shearer*, *Magnets attract iron*, The abbreviation FAQ stands for ‘frequently asked question’, *My niece earns a living as a classically trained musician*).\(^6\) However, the present tense also has secondary meanings, which need not necessarily refer to situations that hold at the present moment. The present tense may also be used to refer to the future (e.g., *The plane leaves tomorrow at 11:15 a.m.*, *I will talk to her as soon as I see her*). In addition, it may be also used to refer to the past (e.g., in newspaper headlines (e.g., *Drunk Hollywood star breaks his leg*) and in narrative descriptions (‘historic present’) (e.g., *I am sitting in her room and all of a sudden this guy walks in and starts flirting with me*). A second commonly discussed aspect of the relationship between tense and time is that the grammatical category of tense does not always express temporal meanings. Tense is often used to express hypothetical meaning as is the case with the past in, for example, conditional sentences (e.g., *If I had children, I would not allow them to talk to me like that*). Moreover, modality may also be expressed by means of the grammatical category of tense (e.g., a will-future in *He will come to tomorrow’s event even if I have to drag him!*). In this doctoral dissertation, the focus will be on the temporal meanings of the grammatical category of tense, more specifically on the temporal meanings of the past and the present perfect when used to locate bygone situations in English and in Dutch.

The two items of agreement discussed above are counterbalanced by a few items of disagreement, which disclose what is referred to by Declerck (1991) as a “lack of consensus” (p. 8) with regard to the concept of tense. The first item of disagreement concerns the precise number of tenses that are said to make up the tense paradigm in a language. Paraphrased in a more language-specific context, that is, an English-specific context, this item could be reduced to one single question: How many tenses does present-day English have at its disposal? The answer to this question varies according to the scope of the definition of tense that is applied. The two answers—which represent extreme answers to

\(^6\) Some uses of the present tense (e.g., general truths/facts/timeless statements such as *Magnets attract iron*) refer to situations that not only hold at the present time but also held at past times and will hold at future times. However, there are scholars who question that such instances of references to past and future times are inherent features of the present tense. They explain such references as interpretations which are the result of implications which are derived from world knowledge and/or from the type of situation which is found in the actual clause (e.g., Downing & Locke, 2006).
this tense-related question—are said to be two tenses (according to an extremely narrow definition of tense) and sixteen tenses (according to a much broader definition of tense) (Kortmann, 1999). However, at the same time, other answers, such as three, six and twelve tenses, would also be perfectly possible depending on the criteria used to define the concept of tense. Let us have a look at both the possibilities available to expand the English tense system and the various lines of logic used to explain the scope of the grammatical category of tense and some of the numbers listed above.

Many linguists claim that English has only two tenses (present and past) and that tense is a grammatical category which is expressed morphologically, more precisely, by means of inflectional morphemes only. Consequently, any verb forms which are not created by means of inflectional tense morphology (e.g., the present perfect, the analytic will-future) are a priori not considered true tense forms (e.g., Finch, 2000; Thornbury, 2006; Trask, 1997, 2000). There are, however, many linguists who do not subscribe to such a narrow definition of tense and who include, for example, the analytic will-future as a possible tense in English (e.g., Crystal, 2003; Declerck, 1991, 2003, 2006), bringing the total amount of tenses to at least three (present, past and future). This is a highly controversial issue but the logic behind such a decision will be explained below in the discussion of Declerck’s taxonomy of tenses in present-day English.

Another option is to broaden the definition of tense even further and to include the perfect verb forms (present perfect, past perfect, future perfect), often referred to as aspeical variants of non-perfect verb forms, which brings the total number of tenses to at least six (e.g., Declerck, 1991, 2003, 2006). Once again, the inclusion of perfect forms in the tense system is not an uncontroversial issue but a possible line of logic behind such a decision will also be provided below. In theory, the number of six tenses could be doubled by including also the progressive verb forms as possible tense forms, thus yielding twelve tenses. And if the conditional verb forms were added to those twelve tense forms, the result would be a total of sixteen tenses in present-day English (see Figures 1.1 and 1.2).

---

7 Many linguists refer to the present tense using the term non-past since they believe that the term non-past more accurately expresses the use of the present tense as a means of referring to times other than the past (Trask, 2000). Indeed, the present tense is often used to refer to times other than the present. However, the present tense is also used sometimes to refer to the past (e.g., in newspaper headlines and in narrative descriptions ("historic present")). Thus, the term non-past itself is not as accurate as some linguists claim that it is. Consequently, the term non-past will not be used in this dissertation unless it is explicitly used in references to its use by other scholars.

8 Many grammars of English list the auxiliary shall as a formal and/or less usual alternative auxiliary for the will-future in the first person (singular and plural) (Carter & McCarthy, 2006; Declerck, 2003; Mackenzie, 1997; Ungerer, 2000). However, because the shall-alternative is not as frequent as the more widespread future tense auxiliary will, the concept of will-future will be used in this dissertation to refer to both instantiations of this future form in English.
The inclusion of progressive forms in a possible tense paradigm stretches the narrow definition of tense quite considerably. Many linguists insist that the progressive verb forms are not separate tense forms but rather aspectual variations of non-progressive verb forms. However, grammars of English which have a strong(er) pedagogical focus often implicitly or explicitly include the progressive verb forms in the tense system in such a way that progressive verb forms are represented as verb forms with tense status in their own right rather than aspectual variations of simple verb forms (e.g., Aitken, 1992; Alexander, 1988; De Moor, 1998; Hewings, 1999; Murphy, 1985; Thomson & Martinet, 1986a, 1986b, 1986c).

In his descriptive theory of tense, Declerck (1991, 2003, 2006) recognizes a total of eight common tenses in present-day English, which he refers to using the following grammatical terminology: (1) present, (2) present perfect, (3) past/preterit(e),9 (4) past perfect, (5) future, (6) future perfect, (7) conditional and (8) conditional perfect. A total of eight common tenses may be a surprise to those who define tense extremely narrowly, so let us have a look at some of Declerck’s guiding principles for drawing up this specific taxonomy of tenses in English. Some of the principles have already been referred to above but will be examined in more detail below. In addition, possible controversial issues which

---

were introduced above in the discussion of definitional items of disagreement will also be explained.

![Figure 1.2. Conjugation of the verb talk in a possible 16-tense verb system in present-day English](image)

The first controversial issue is the listing of compound verb forms as possible tensed verb forms (in present-day English) based on morphological features of the verb phrase. Let us investigate this issue by discussing the nature of future verb forms in present-day English. In his list of common tenses in English, Declerck refers to two future tenses in English: the future (e.g., *I will talk to him next time*) and the future perfect (e.g., *I will have talked to him by then*). Two main problems are usually addressed in discussions of the status of future tenses in present-day English: (1) the analytic nature of future verb forms and (2) the modal features of the future auxiliary *will*. For the remainder of the discussion, the second problem will not be of much interest. Consequently, it will not be discussed in great detail. The first problem, the analytic nature of future verb forms, which was already briefly mentioned above, concerns the fact that present-day English does not generally use synthetic verb forms to refer to the future but, instead, uses a range of options to refer to the future, most of which are morphologically analytic.\(^\text{10}\)

The most neutral way of referring to the future

---

\(^{10}\) Trask (1997) defines the term *analytic* as “a label applied to a grammatical form which is constructed by adding additional words to the word being inflected” (p. 14) and the term *synthetic* as “a label applied to a grammatical form which is constructed entirely by affixing or modifying the
in present-day English is by means of the analytic will-future (Kortmann, 1999). Many linguists do not regard the English will-future as a tense form in its own right since it is formed using the future auxiliary will in combination with the base form of the main verb and not by means of any inflectional tense morphology. In other words, those who do not regard the English will-future as a tense define the grammatical category of tense extremely narrowly. According to this logic, any verb form which is not created using only inflectional tense morphology is not considered a true tense (form). The creation of future tense forms by means of non-inflectional tense morphology (i.e., by means of analytic verb forms) is not unusual in language and it is a linguistic feature in English and in, for example, other Germanic languages such as Danish, Dutch, German, Icelandic, Norwegian and Swedish. Such instances of non-inflectional future tense morphology contrast with the most common future tense forms in, for example, Romance languages such as French, Italian, Portuguese and Spanish, which do generally have synthetic future verb forms at their disposal—often in addition to other verb forms with future meanings—and which use these synthetic verb forms to refer to the future (See Table 1.1).

Declerck recognizes the distribution of analytic and synthetic future verb forms referred to above but does not agree with the assumption that tense should be conceived of as a grammatical category which is marked morphologically by means of inflectional morphemes only. “It is well known”, Declerck (1991) argues, “that it is often the case that one and the same idea is expressed morphologically in one language and by means of an independent morpheme in another” (pp. 9–10). He exemplifies this statement by making a passing reference to the morphological features of the definite article in Swedish. In English, the definite article is always realized as a free, lexical morpheme, which is placed in front of the singular or plural noun which it determines. However, in Swedish the definite article is often—though not always—realized as a single suffix, which is generally added to the singular or plural noun which it determines (e.g., bil/bilen (car/the car), bilar/bilarna (cars/the cars), äpple/äpplet (apple/the apple), äpplen/äpplet (apples/the apples). This feature, which is also referred to as the suffixed article, is by no means exclusive to Swedish and is also found—albeit in varying forms of distribution—in other languages (e.g., Albanian, Bulgarian, word in question, without the use of any additional words” (p. 215). With respect to verb forms the analytic/synthetic distinction is reflected in, for example, the following forms: he will eat (analytic, because of the future auxiliary will), he eats (synthetic, because of the inflectional suffix -s).

11 It should be stressed that both Germanic languages and Romance language have other (non-future) verb forms at their disposal to refer to the future but for the discussion at hand this feature is irrelevant and will consequently not be discussed. The typological distinction between analytic and synthetic languages and language forms is not a sharp one. Instead, the distinction should be seen as a continuum with most languages located at various points between the two extremes of fully analytic and fully synthetic (Barber, 2000).
Danish, Icelandic, Norwegian, Romanian).\textsuperscript{12} With reference to tense, Declerck (1991) subsequently poses the following question: “Why should we not accept that some tenses can be expressed morphologically while others make use of free morphemes (auxiliaries)?” (p. 10). Although Declerck’s reasoning may be viewed by some linguists as simplistic, it does bring home an issue which is found not only in the discussion of tense: the fact that languages may express the same grammatical feature (e.g., aspect, definiteness, number, tense) but that they need not necessarily select the same resources to express these features.

<table>
<thead>
<tr>
<th>Table 1.1. Distribution of analytic and synthetic future verb forms in Germanic and Romance languages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Germanic languages</strong></td>
</tr>
<tr>
<td>Danish</td>
</tr>
<tr>
<td>Dutch</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>German</td>
</tr>
<tr>
<td>Icelandic</td>
</tr>
<tr>
<td>Norwegian</td>
</tr>
<tr>
<td>Swedish</td>
</tr>
<tr>
<td><strong>Romance languages</strong></td>
</tr>
<tr>
<td>French</td>
</tr>
<tr>
<td>Italian</td>
</tr>
<tr>
<td>Portuguese</td>
</tr>
<tr>
<td>Spanish</td>
</tr>
</tbody>
</table>

As described above, the analytic nature of the future verb forms in present-day English is not the only problem that one encounters in discussions of the tense status of future verb forms in English. In addition to the morphology of future verb forms, many linguists do not recognize any future tense(s) in present-day English on the assumption that the semantics of the will-future verb forms are an example of the fact that the future “is not a tense at all, but a mode” (Cygan, 1972, p. 9, quoted in Declerck, 2006, p. 102). Declerck (2006) recognizes that the future tense may have modal aspects of meaning (e.g., subjectivity) and that there is always an element of epistemic modality in the meaning of a future tense. However, he also claims that some uses of the will-future are close to

\textsuperscript{12} The rules regarding the forms, meanings and uses of the definite article in Swedish are more complex than computationally adding the appropriate suffixes to singular and plural forms of nouns. However, for the discussion at hand such nuances are irrelevant and will not be discussed. For more detailed information on this feature of Swedish grammar see, for example, Holmes and Hinchliffe, 1994, Meijer, 1997, Ramge, 2002.
being purely temporal since they do not involve any modal aspect of meaning. Consequently, Declerck does recognize future tenses in English.

The second controversial issue concerns the listing of a small group of verb forms as possible examples of tense forms (in present-day English) which are actually referred to as aspectual verb forms in many descriptions of English verb forms. The perfect verb forms fall into this category of verb forms. Declerck distinguishes four perfect tenses in present-day English: (1) the present perfect (e.g., She has studied), (2) the past perfect (e.g., She had studied), (3) the future perfect (e.g., She will have studied) and (4) the conditional perfect (e.g., She would have studied). Like future verb forms, perfect verb forms are ‘at a morphological disadvantage’ for acceptance as tenses as a result of their analytic nature, that is, they are formed using a (tensed) form of the auxiliary have and the past participle of the main verb. In addition, many linguists do not consider perfect verb forms as tense forms in their own right but rather as aspectual variations of existing (non-perfect) verb forms. According to one group of linguists, the present perfect can be considered a combination of a present tense and an aspectual meaning component of current relevance (Declerck, 2006). Another group of linguists claims that the present perfect is actually a representation of the same temporal structure expressed by the past but that the present perfect “differs from the past only in that it also expresses ‘perfect aspect’” (Declerck, 2006, p. 109). Declerck, however, does not adhere to these views and attributes a tense structure to the present perfect which is different from the tense structures found for either the present tense or the past tense.

The features that have been discussed above show that reaching agreement on a possible definition of the grammatical category of tense is a problematic endeavour. Although there are some commonly accepted definitional features which are found in many existing definitions of tense, there are also several controversial features, which render any attempt at defining tense problematic. Two of the most controversial issues that are generally addressed in definitions of tense are the morphological status (analytic versus synthetic) and the aspectual features of verb-phrase elements. For both of these controversial issues arguments may be found in favour of or against accepting specific verb forms (e.g., perfect verb forms) as tenses in their own right (see discussion above). The approach to tense that will be adopted in this doctoral dissertation is the one put forward by Declerck (1991, 2003, 2006). The most updated and most

---

13 The concept of epistemic modality is defined by Declerck (2006) as the “modality having to do with the possible degrees of the speaker’s commitment to the truth of a proposition” (p. 775). According to Declerck (2006) such modality may be expressed by “modal adverbs like certainly, perhaps, possibly, etc. or by auxiliaries like must, should, ought to, will, can, could and need” (p. 775).

14 The aspectual meaning of current relevance in perfect verb forms is also described using the adjective resultative (Leech, 2004). However, the idea of current relevance is not completely unproblematic (Lewis, 1986). As a result, it will not be used in the further discussion of perfect verb forms and of the present perfect tense unless it is explicitly used in the literature on perfect tenses.
comprehensive version of Declerck’s theory of tense will be introduced in the following section. It is the concepts that are put forward below which will be used in the further development of ideas in this chapter.

1.3 Conceptualizing tense

1.3.1 Linguistic time in English

As was mentioned in Section 1.2, Declerck (2006) defines tense as “a linguistic concept (as opposed to time) denoting the form taken by the verb to locate (the actualization of) the situation referred to in time” (p. 820). The concept of time, however, is not as straightforward as one might expect and this has been reflected in Declerck’s work and in the overall development of his descriptive theory of tense in English. Declerck makes clear references to, for example, the difference between objective (physical) time and linguistic time (Declerck, 1991, 2003, 2006). Whereas objective (physical) time consists of two parts (the past and the future), which are separated by the present, linguistic time, according to Declerck (1991), is “time as it is perceived and talked about by language users” (p. 16).

English-speaking language users generally divide linguistic time into two time-spheres: (1) the present time-sphere and (2) the past time-sphere (Declerck, 1991, 2003, 2006). Characteristic of these two time-spheres is that the past time-sphere is defined as an (indefinite) time span which is located entirely before the temporal zero-point ($t_0$) and which does not include $t_0$ (Declerck, 1991, 2006). Declerck (2006) defines the temporal zero-point as follows:

> The time which is the ultimate ‘origin’ of all the temporal relations expressed by the temporal structure of a tense, i.e. the only time in a tense structure that is not itself represented as dependent on another (more basic) time. It is the only time that is given (‘assumed known’) whenever a sentence is uttered. In English, the temporal zero-point is nearly always the encoding time, i.e. the time of uttering or writing the message. Occasionally, the zero-point is the decoding time, i.e. the time when the addressee is expected to hear or read the message, as is the case when a note stuck to someone’s door reads *I am in room 21*. (As always, the present tense locates the situation time at $t_0$, but $t_0$ is the time of reading the message rather than the time of writing it.) (p. 820, quoted with original highlighting)

The present time-sphere, on the other hand, is defined as an (indefinite) time span which does include $t_0$. The presence of $t_0$ in the present time-sphere leads to the division of that time-sphere into three zones (or sectors): (1) the pre-present zone (i.e., the part of the present time-sphere which precedes $t_0$), (2) the present zone (i.e., the part of the present time-sphere which coincides with $t_0$) and (3) the post-present zone (i.e., the part of the present time-sphere which

---

15 A time-sphere refers to a length of time. By no means should this concept be interpreted as spherical (Declerck, 2006).

16 The temporal zero-point is also referred to as the time of utterance (TU) (Declerck, 2003).
follows to) (Declerck, 1991, 2003, 2006). The two basic divisions referred to above are schematized in Figure 1.3.

![Diagram](image)

Figure 1.3. Linguistic conceptualization of the time line in English (adapted from Declerck, 2006)

The eight common English tenses referred to above are used to locate situations in either the present time-sphere or past time-sphere. Consequently, Declerck divides the English tenses into two main categories: (1) present (time-sphere) tenses and (2) past (time-sphere) tenses. The category of present (time-sphere) tenses comprises the present tense, the present perfect, the future tense and the future perfect tense and the category of past (time-sphere) tenses comprises the past tense, the past perfect, the conditional tense and the conditional perfect.\(^{17}\)

The two divisions above (i.e., the division into time-spheres and the division into zones/sectors) form the basic, temporal/conceptual framework in which the English tenses may be used to locate situations. However, when considering the meanings and uses of tenses several factors must be considered. One important factor, if not the most important factor, are the semantics of the individual tenses. Declerck (2006) recognizes this and refers to the semantics issue as follows: “The use of a tense is wholly determined by its semantics (= temporal structure), which has to fit in with the temporal information given by the time-specifying adverbial or by the context” (p. 599). This brings us to an aspect of tense which is considered extremely challenging by linguists, L2 teachers and L2 learners: Even though tense can be found at a subsentential level and at a sentential level (i.e., in simple and complex clauses), it is often active also at a suprasentential level (i.e., in discourse). At both the sentential level and

---

\(^{17}\) In theory, the eight common tenses recognized by Declerck should actually be interpreted as umbrella terms since each tense comprises both the simple form and the progressive form of that tense. This means that, for example, the present tense comprises the simple present (e.g., I talk) and the present progressive (e.g., I am talking). This principle applies to all of the eight common tenses recognized by Declerck.
the suprasentential level, situations in clauses are often temporally related to each other by means of temporal domains (Declerck, 2006). The concept of temporal domains can be exemplified using the following sentence:

(1.1) The student told the teacher that he *had handed in* the wrong paper and that he *would hand in* the right paper on Tuesday.

In (1.1), the reference is clearly to a past domain, that is, to a structured set of times, which is located in the past time-sphere and which consists of the times of the three situations which are referred to in the sentence (*told* (past), *had handed in* (past perfect) and *would hand in* (conditional)). The first situation (*told* (past)) is the situation which establishes the past domain and which is therefore referred to as the central situation. A tense that is used to refer to such a central situation is called an absolute tense.\(^{18}\) The other two tense forms in the sentence above (*had handed in* (past perfect), *would hand in* (conditional)) both refer to situations which are temporally subordinated (= temporally related) to the central situation: the situation referred to by *had handed in* is represented as anterior to the situation of John telling the teacher about what had happened whereas the situation referred to by *would hand in* is represented as posterior to the situation of John telling the teacher about what had happened. A tense that is used to refer to such temporally subordinated situations is called a relative tense.

Within Declerck’s descriptive theory of tense, the English-speaking language user’s task (in discourse) consists of having to decide, for every new clause, whether the situation contained in that clause should be incorporated into the existing temporal domain, that is, whether the existing temporal domain should be expanded, or whether a new temporal domain should be established, into which further situations may subsequently be incorporated. Establishing a new temporal domain is referred to by Declerck (1991, 2003, 2006) as a shift of domain. If language users decide to incorporate a new situation into an existing temporal domain, they are said to resort to temporal subordination, which involves temporally binding the new situation to an orientation time other than the temporal zero-point (Declerck, 1991, 2003, 2006).

### 1.3.2 Relative tense: The expression of tense relations within a domain

In addition to dividing linguistic time into two time-spheres and the present time-sphere into three zones, English-speaking language users also use various language-specific systems to express tense relations (T-relations) in a domain established in the past or in the present time-sphere. In essence, there are three types of (domain-internal) relations that can be expressed linguistically by means of tenses: (1) anteriority (e.g., *When I arrived at the party, she had already left*),

\(^{18}\) English has a total of four absolute tenses: (1) the past, (2) the present perfect, (3) the present and (4) the future (Declerck, 2006).
(2) simultaneity (e.g., *That man said that he was looking for his partner*) and (3) posteriority (e.g., *They promised her that they would visit her every week*). The relevance of these domain-internal relations in the language acquisition/teaching process will be discussed in the next section, which will shed additional light on the T-relations of the temporal form–meaning–use (FMU) mappings under investigation and will present the temporal FMU mappings under investigation by placing them into Declerck’s descriptive theory of tense. In addition, a crosslinguistic comparison with Dutch will be drawn up to highlight similarities and differences between these typologically closely related languages.

1.4  **Locating bygone situations: The past and present perfect from a crosslinguistic perspective**

The aim of this section is twofold: (1) to introduce the past and the present perfect, which may be used in both English and Dutch to locate bygone situations, and (2) to provide more detailed information on the nature of these temporal FMU mappings in English and in Dutch, which are typologically closely related languages. Since this section simply seeks to introduce Declerck’s linguistic analysis of the past/present distinction, the information provided will be limited—as much as possible—to a purely (cross)linguistic description of the FMU mappings under investigation.\(^{19}\) This section will first highlight how the past and the present perfect fit into Declerck’s descriptive theory of tense in English introduced above. In addition, the important interplay between these temporal FMU mappings and their linguistic environments (e.g., adverbials) in English will also described. Subsequently, a crosslinguistic comparison between the English FMU mappings and their Dutch counterparts will be drawn up to highlight the most important similarities and the differences between the mappings in both languages.

1.4.1  **The past and present perfect in English**

In his descriptive theory of tense, Declerck recognizes the problem involved in choosing between the past and the present perfect. He narrows this problem down to the context of using either the past or the present perfect to refer to what he calls bygone situations (Declerck, 1991, 2003, 2006). A bygone situation is defined by Declerck (1991, 2003, 2006) as a situation which precedes the temporal zero-point \(t_0\). If this concept is schematized using Declerck’s linguistic conceptualization of the time line in English presented in Section 1.3.1, a bygone situation is located in either the past time-sphere or the pre-present zone (in the present time-sphere). Therefore, bygone situations may be temporally realized using either the past or the present perfect (see Figure 1.4).

---

\(^{19}\) For any complexity-related and SLA-related issues dealing with the temporal FMU mappings discussed in this section see Chapters 2 and 3 respectively.
In discussing the nature of this temporal choice, Declerck (2006) makes a distinction between the use of the past and the present perfect in clauses without temporal adverbials and the use of these two tenses in clauses with temporal adverbials. In both cases, language users—or speakers in Declerck’s terminology—are often ‘forced’ to choose between the past and the present perfect. Consequently, both scenarios will be investigated here.

The discussion below reflects the distinction between clauses with and without temporal adverbials, and will highlight the issues relevant to the meanings and uses of these tenses in past-zone contexts.

**Clauses without temporal adverbials.** In clauses without temporal adverbials, Declerck’s (2006) general description regarding the use of the past and the present perfect is as follows:

> The speaker is concerned with NOW when he uses the present perfect and with THEN when he uses the preterite. In other words, the present perfect implies that the speaker’s ‘temporal focus’ is on the pre-present zone, whereas the past tense puts the temporal focus on the past time-sphere. Whether the speaker is concerned with NOW or THEN can be clear from the way the sentence is used in context or from particular constituents of the sentence. (p. 317, quoted with original highlighting)

Although the description above is relatively succinct, its scope is extremely broad in that Declerck uses it to cover almost all cases in which a choice has to be made between the past and the present perfect to locate bygone situations in clauses without temporal adverbials (and, as a matter of fact, also in clauses with temporal adverbials, which will be discussed later on). Questions, however, remain: What principles are in force in English to guide speakers in their

---

20 The terms *language user* and *speaker* will both be used in this dissertation. Although Declerck (1991, 2003, 2006) uses the noun *speaker* to refer to a language user, he does not limit his descriptive theory of tense to the oral mode.

21 The reader is advised to bear in mind that both the past and the present perfect may also be used in English to refer to non-bygone situations. For the current discussion, however, these uses are of no interest and will not be featured in this chapter.
When referring to bygone situations, (non-interrogative) clauses without any time-specifying adverbials generally allow the use of both the past and the present perfect (Declerck, 2006). The past is used when the speaker is somehow concerned with THEN (e.g., when the speaker is thinking of the time of actualization of the bygone situation or is expressing concern with the actualization of the situation itself (e.g., where?, why?, how?)). In contrast, the present perfect is used when the speaker is somehow concerned with NOW (e.g., when the speaker is announcing the actualization of the bygone situation as news or is expressing concern with the current relevance of the bygone situation). If this idea is linked to the concepts of temporal focus and the time line in English above, the past places the temporal focus on the past time-sphere and the present perfect places the temporal focus on the present time-sphere (see Figure 1.4). Declerck uses the idea of actualization focus (i.e., concern with some aspect of the bygone situation (where?, why?, how?)) to explain the use of the past in clauses without time-specifying adverbials. In addition, he discusses the idea of introducing a new topic into the current discourse (i.e., current relevance) to explain the use of the present perfect in clauses without time-specifying adverbials. Declerck (2006) continues his discussion of this specific feature by providing clear examples and fine-tuning his theory.

Even though the speaker may choose to focus on NOW or THEN when locating bygone situations in clauses without temporal adverbials, there are certain principles with respect to the use of these two tenses, which Declerck describes in his theory of tense. All the instances of past tense use may be subsumed under the feature of actualization focus (Declerck, 2006). If the speaker is concerned with the actualization of the bygone situation, a past will be chosen to reflect this concern. This use of actualization focus is shown in the following examples taken from Declerck’s (2006) work:

---

Actualization focus may be realized in two ways: (1) focus on the actualization of the bygone situation as a whole or (2) focus on a specific aspect of the bygone situation (e.g., where, why or how the bygone situation took place) (Declerck, 2006). It should be stressed, however, that question words such as where, why, how do not automatically exclude the use of a present perfect. In discourse, there are three conditions in which wh-questions may contain present perfect forms if the environment for using the present perfect is available: (1) if the meaning to be expressed is a continuative meaning (e.g., Why have you been seeing her behind my back?), (2) if the meaning to be expressed is one of an indefinite perfect (e.g., Why has he never tried to talk about this issue with me?) and (3) if the speaker is focusing on the present result (i.e., the (current) relevance) of the bygone situation (e.g., Where has she put my keys?, the speaker is interested in knowing where the keys are now) (Declerck, 2006).

---
Sentences (1.2) and (1.3) show how speakers may choose between either the past or the present perfect depending on whether they wish to make use of actualization focus or not. However, there are instances where speakers are more restricted and are ‘forced’ to use one of the two tenses to express specific meanings. When-questions and clauses containing definite noun phrases entailing actualization focus are two such examples. Consider the following examples:

(1.4) When did I swear at you?
(1.5) Did you buy her those expensive shoes?

If speakers use a when-question to enquire about the actualization of a bygone situation, they take for granted that the situation has actualized. Consequently, they are ‘forced’ to use a past. However, this does not mean that a present perfect in when-questions is impossible (e.g., When have I sworn at you?). By using a present perfect, the speaker does not express an interest in finding out when the situation actualized but rather an interest in asking the addressee whether an occasion (or occasions) on which the situation actualized can be provided. Similarly, when a clause without temporal adverbials contains a definite noun phrase entailing actualization focus, the past must be used (see sentence (1.5) above). The presence of the definite noun phrase those expensive shoes entails that the bygone situation being referred to (i.e., the buying of the shoes) is assumed to be identifiable. In other words, the definite noun phrase entails actualization focus.

**Clauses with temporal adverbials.** Clauses with temporal adverbials pose a different challenge in that compatibility (i.e., semantic alignment) must be achieved between the use of tenses, on the one hand, and the temporal adverbials on the other hand. Various classes of temporal adverbials may be distinguished and since class membership may influence the choice between the past

---

23 The use of temporal adverbials is not limited to the sentence in which the tense appears. Often, they may also be found at a suprasentential level (e.g., in extended discourse) and contribute to the overall temporal features of a longer stretch of text. See Chapters 3, 4, 5 and 6 for examples of suprasentential features of temporality.
and the present perfect, a more detailed look at the classes of temporal adverbials is warranted at this point.

In his descriptive theory of tense, Declerck (2006) divides temporal adverbials into deictic (anchored) or non-deictic (unanchored) temporal adverbials, depending on whether the adverbials are related to a temporal anchor (t₀ or another time) or not. Examples of deictic temporal adverbials are *last night, the next day, today, tomorrow, yesterday, two days ago, two days later*. Examples of non-deictic temporal adverbials are *at seven o’clock, every year, in 2007, on Mondays, some time or other*. Consider the following sententially contextualized examples:

(1.6) I saw her yesterday. (deictic adverbial)

(1.7) The woman showed up two days later. (deictic adverbial)

(1.8) We generally have dinner at seven o’clock. (non-deictic adverbial)

(1.9) I met Madonna in 2007. (non-deictic adverbial)

In turn, deictic temporal adverbials can be further categorized into absolute deictics and relative deictics. Absolute deictics are temporal adverbials that are anchored to t₀ (e.g., *this morning, three weeks ago, today, tomorrow, yesterday*) and relative deictics are temporal adverbials that are anchored to another time and not to t₀ (e.g., *the next day, the same day, two days earlier*).

(1.10) The e-mail that I sent to you yesterday was sent back to me. (absolute deictic temporal adverbial)

(1.11) The e-mail that I had sent to you two days earlier was sent back to me. (relative deictic temporal adverbial)

Important for the analysis of L2 learner use of the past and the present perfect is Declerck’s (2006) categorization of deictic temporal adverbials into single-zone adverbials and multi-zone adverbials. Single-zone adverbials refer to only one absolute zone (e.g., past-zone adverbials refer to only the past zone) whereas multi-zone adverbials refer to a timespan which may include more than one time zone (e.g., *today, which may include the present zone as well as the post-present, the pre-present or the past zone depending on the choice of tense*). Consequently, multi-zone adverbials are compatible with more than one absolute tense.

(1.12) I spoke to him yesterday. (single-zone adverbial)

(1.13) John is in London today. (multi-zone adverbial, temporal focus on present)

(1.14) John has been in London today. (multi-zone adverbial, temporal focus on some pre-present part of today)
(1.15) *John was in London today.* (multi-zone adverbial, temporal focus on some past time in the course of today)

(1.16) *John will be in London today.* (multi-zone adverbial, temporal focus on some post-present time in the course of today)

In addition to single-zone and multi-zone adverbials, English also has zone-independent adverbials (Declerck, 2006). Zone-independent adverbials are adverbials that specify a time which is not linked to one specific zone (e.g., *at seven o’clock, every year, on Mondays*). In terms of compatibility, zone-independent adverbials are also compatible with various absolute tenses, and choosing the absolute tense is dependent on the temporal focus. Zone-independent adverbials are inherently non-deictic (Declerck, 2006).

Declerck makes several further distinctions in his categorization of time-specifying adverbials (e.g., homogeneous versus heterogeneous time-specifying adverbials, inclusive versus non-inclusive heterogeneous time-specifying adverbials) but since these distinctions are not relevant to the further discussion of the temporal FMU mappings being investigated, they will not be discussed here. A summary of Declerck’s categorization of time-specifying adverbials into single-zone, multi-zone and zone-independent adverbials may be found in Table 1.2.

The class of past-zone adverbials is of special interest for the discussion of the temporal FMU mappings under investigation in this doctoral dissertation. Declerck (2006) defines a past-zone adverbial as a “single-zone time-specifying adverbial which specifies a time in the past zone” (p. 798). As far as the grammatical category of tense is concerned, past-zone adverbials can be used in combination with the past but under no circumstances can they be used with the present perfect in grammatically correct standard present-day English.24 In other words, once a past-zone context has been established by means of, for example, one or several implicit or explicit time-specifying adverbials, grammatically correct standard present-day English does not generally allow the use of the present perfect (see example sentences 1.17–1.32).

---

24 The concept of standard present-day English is an umbrella term and is used here to refer to the two most influential regional varieties of English: British English (BrE) and American English (AmE). However, two issues should be mentioned in this respect. Firstly, the selection of BrE and AmE by no means reflects any negative evaluation of other regional varieties of English (e.g., Australian English, Canadian English, South African English). BrE and AmE are simply the two most widely taught standard forms of English worldwide. Secondly, the two most influential regional varieties of English are themselves simplifications of a more complex picture since both BrE and AmE are, in fact, characterized by dialect variation resulting in differences at various language levels (e.g., lexical, phonological, morphological, syntactic) including, at times, differences in verb usage (Leech, 2004).
Table 1.2. Zone-specific categorization of time-specifying adverbials

- **deictic (anchored) adverbials**

  **single-zone adverbials**
  (a) past-zone adverbials (e.g., last year, two days ago, yesterday)
  (b) pre-present-zone adverbials (e.g., since then, so far, up to now)
  (c) present-zone adverbials (e.g., at this moment, right now)
  (d) post-present-zone adverbials (e.g., next week, tomorrow)

  **multi-zone adverbials** (e.g., always, forever, today)

- **non-deictic (unanchored) adverbials**

  **zone-independent adverbials** (e.g., at seven o’clock, on Mondays)

### 1.4.2 The past tense and present perfect in Dutch

Declerck’s (1991, 2003, 2006) descriptive theory of tense is one which has been developed with English in mind as the main object of investigation. Consequently, Declerck’s theory as a whole is unique to the English tense system. Although individual features in the theory are shared by other languages, the specific constellation of the individual features in Declerck’s theory of tense is what makes up the entire system which is referred to as the English tense system. This section, however, will place the focus on the Dutch tense system. As such, this section has two aims. Firstly, to show that there are indeed specific features of Declerck’s theory of tense—more specifically features related to the past and the present perfect when used to locate bygone situations in English—that are also found in Dutch. Secondly, to show that, for all the shared features, there are also fundamental differences which must be considered when both English and Dutch come into contact with each other (e.g., in the Dutch-speaking ESL learner’s mind). The features which were highlighted in the discussion of the past and the present perfect in English will be used as anchor points but since there is no descriptive theory of tense in Dutch as comprehensive as Declerck’s theory of tense in English, a collection of other sources had to be consulted. In total, three main types of source material will be used in this section: (1) descriptions of the Dutch tense system generally meant for native speakers of Dutch (e.g., Haeseryn, Romijn, Geerts, de Rooij, & van den Toorn, 1997), (2) descriptions of the Dutch tense system meant for English-speaking L2 learners of Dutch (e.g., Donaldson, 2008; Fehringer, 1999; Shetter & Ham, 2007) and (3) descriptions of the Dutch and English tense systems found in contrastive grammars of English and Dutch (e.g., Aarts & Wekker, 1993; Mackenzie, 1997; van Brederode & Koopman, 1990). In addition, a fourth type of source material will also be featured but only sporadically: English grammars
meant for L2 learners of English who are not Dutch-speaking. This may seem somewhat contradictory since the aim of this section is to shed light on the Dutch tense system and on possible similarities and differences between the Dutch tense system and the English system. However, the similarities and differences between the English and Dutch tense systems are not unique to this specific language pair. The Dutch tense system shares many similarities with the tense systems of other typologically closely related languages (e.g., French, German). These similarities may be grouped to highlight fundamental differences with the English tense system. The language which will also be considered here is German since it is also a Germanic language and of all the Germanic languages is most closely related to Dutch. Consequently, several English grammars meant for German-speaking L2 learners (e.g., Hoffmann & Hoffmann, 2001, 2005; Ungerer, 2000; Ungerer, Meier, Schäfer, & Leechler, 2009) will also be used to draw comparisons and as resources of authentic example material.

Before discussing the Dutch tense system in more detail, one terminological issue must be addressed. In Dutch, a similar temporal choice is made to locate bygone situations. Whereas in English the choice is one between the past and the present perfect, in Dutch the choice is one between two tenses which are generally referred to in Dutch grammars as the onvoltooid verleden tijd (ovt) (‘uncompleted past tense’) and the voltooid tegenwoordige tijd (vtt) (‘completed present tense’). In many Dutch grammars, the Latin names for these tenses are also used: imperfectum for the onvoltooid verleden tijd and perfectum for the voltooid tegenwoordige tijd. As far as form, meaning and use are concerned, the ovt resembles the English past (i.e., past of the main verb) and the vtt resembles the English present perfect (i.e., the present of an auxiliary verb with the past participle of the main verb) even though differences may be found at all three levels of the mappings. Consequently, the following discussion will use the terms Dutch past (tense) and Dutch present perfect (tense). The use of these terms focuses on the similarities between the English tenses and the Dutch tenses. However, differences between the English tenses and the Dutch tenses should not be overlooked.

The most comprehensive grammar of Dutch for native speakers of Dutch is Haeseryn et al.’s (1997) Algemene Nederlandse Spraakkunst. In the chapter on the functions of the Dutch tenses, Haeseryn et al. (1997) provide a detailed description of the various functions of the tenses in Dutch, supplying not only theoretical explanations but also practical examples. The approach that they adopt is one which relies heavily on existing analyses of tense in language in general and especially on Reichenbach’s influential analysis found in Elements of Symbolic Logic (1947). Reichenbach’s analysis of tense uses three elements to describe tenses: (1) a point of the event (E), (2) a point of reference (R) and (3) a point of speech (S). Haeseryn et al. (1997) also use three elements in their analysis and provide the following Dutch names for their elements: (1) the werking (W) (the equivalent of Reichenbach’s E), (2) the referentiepunt (R) (the equivalent of
Reichenbach’s R and (3) the spreekmoment (S) (the equivalent of Reichenbach’s S). The Dutch past and the Dutch present perfect are schematized by Haeseryn et al. (1997) as follows:

![Figure 1.5. The representation of the Dutch past using the descriptive elements point of the event, point of reference and point of speech (adapted from Haeseryn et al., 1997, p. 115)](image1)

![Figure 1.6. The representation of the Dutch present perfect using the descriptive elements point of the event, point of reference and point of speech (adapted from Haeseryn et al., 1997, p. 114)](image2)

Although there are many details to Haeseryn et al.’s analysis of tense which are irrelevant for the discussion at hand, there are striking similarities with Declerck’s approach to the placement of the past and the present perfect on the time line in English to locate bygone situations. The most striking are the following:

1. Declerck acknowledges the existence of situations, which Haeseryn et al. call events;

---

25 Haeseryn et al.’s werking, referentiepunt and spreekmoment are synonymous with Reichenbach’s point of the event, point of reference and point of speech respectively (Haeseryn et al., 1997; Reichenbach, 1947). Consequently, Reichenbach’s abbreviations will be used since they refer to the English equivalents of Haeseryn et al.’s Dutch terms and since the further discussion of tense in Dutch will be conducted in English. The only exception is Haeseryn et al.’s werking, which will be referred to as event (or the plural form events). The reason for this decision is that the Reichenbachian analysis systematically uses the concept of points, whereas Haeseryn et al.’s analysis does so too but only for point of reference and point of speech.

26 By extension, the similarities are also clear with Reichenbach’s analysis since Haeseryn et al.’s analysis relies heavily on Reichenbach’s analysis. As striking as the similarities between Reichenbach’s and Declerck’s analyses may be, Declerck does express some fundamental objections to Reichenbach’s system of analysing tense. For more detailed information about these objections, see Declerck, 1991.
2. In Declerck’s discussion of the use of the English past and the English present perfect to locate bygone situations, Declerck defines bygone situations as situations which precede the temporal zero-point \( t_0 \). In Haeseryn et al.’s analysis of the Dutch past and the Dutch present perfect, the events are located before the point of speech, which is Haeseryn et al.’s equivalent of Declerck’s temporal zero-point. This location of Haeseryn et al.’s events would qualify them as bygone situations;

3. In Declerck’s analysis of tense in English, the main difference between the use of the past and the present perfect to locate bygone situations is that the past is used when the speaker is somehow concerned with THEN and the present perfect is used when the speaker is somehow concerned with NOW. This concern with either THEN or NOW is reflected in Haeseryn et al.’s analysis by means of the placement of the point of reference. In the case of the Dutch past, the point of reference is represented as being concomitant with the event, that is, with THEN. In the case of the Dutch present perfect, the point of reference is represented as being concomitant with the point of speech, that is, with NOW.

In their discussion of the functions of the Dutch past and the Dutch present perfect, Haeseryn et al. (1997) contrastively elaborate on the uses of both tenses to locate bygone situations. However, they are extremely careful and highlight—at the very beginning of their discussion—that partial overlaps in certain uses of both tenses do exist and that, consequently, differences may be extremely subtle and the degree of acceptability of these tenses may vary from language user to language user. Although Haeseryn et al.’s discussion is a relatively elaborate one considering the material available about the Dutch tense system, it does not contrast the past and the present perfect in English and Dutch. For such a discussion, one is required to look at descriptions of the Dutch tense system drawn up for English-speaking L2 learners of Dutch and descriptions of the English and Dutch tense systems available in contrastive grammars of English and Dutch.

In his comprehensive grammar of Dutch, Donaldson (2008) states the following when starting his discussion of the Dutch present perfect: “It is in the use of the perfect that Dutch tenses differ most from those of English” (p. 187). Donaldson (2008) continues his discussion by providing the following general contrastive guidelines for using the present perfect tense in Dutch:

> All perfects in English are rendered by perfects in Dutch (with one exception, see Present Tense, 11.2.1.3), but most imperfects in English may be rendered by either imperfects or perfects in Dutch, the perfect tense being more common, particularly in speech (for the few instances where English imperfects must be imperfects in Dutch, see Imperfect Tense above). (p. 187)
The example sentences that Donaldson subsequently discusses are the following English sentences:

(1.17) He bought a computer yesterday.

(1.18) He has bought a computer.

In English, a past must be used in (1.17) because of the adverbial yesterday, which is a single-zone adverbial in Declerck’s terminology of adverbials, more specifically a past-zone adverbial. Standard present-day English does not allow the use of the present perfect in past-zone contexts. Dutch, however, often does and translating (1.17) into Dutch would lead, according to Donaldson (2008), to two possible translations:

(1.19) Hij heeft gisteren een computer gekocht. (with a Dutch present perfect)

(1.20) Hij kocht gisteren een computer. (with a Dutch past)

In Donaldson’s (2008) discussion, (1.18) uses a present perfect in English to locate the bygone situation and, in so doing, expresses the language user’s concern with NOW (Declerck, 2006). The Dutch translation would be **Hij heeft een computer gekocht** (with a Dutch present perfect) because of the same temporal logic that applies to the English sentence. Similar contrastive discussions of the past and the present perfect may be found in other publications (Aarts & Wekker, 1993; Fehringer, 1999; Mackenzie, 1997; Shetter & Ham, 2007; van Brederode & Koopman, 1990). Summarizing the statements above, it can be said that both English and Dutch use the past and the present perfect to locate bygone situations. When no implicit or explicit contextual temporal information is provided (e.g., temporal adverbials, world knowledge), language users in both English and Dutch may use the past when they are concerned with THEN and the present perfect with they are concerned with NOW. However, as soon as contextual temporal information becomes retrievable—either implicitly or explicitly—language users must ensure that semantic alignment is achieved and that the selected tense fits in with the temporal structure (i.e., the semantics) of the context. This is where English and Dutch may differ in certain respects when a choice has to be made between the past and the present perfect to locate bygone situations. English does not allow the use of the present perfect in past-zone contexts whereas Dutch often does and even prefers the present perfect to possible past alternatives. However, often any preferences in Dutch do not appear to be guided by a set of clear guidelines (see, for example, Haeseryn et al., 1997). Example sentences (1.17) through to (1.32) show how the similarities and differences between the English and Dutch tense systems are generally highlighted and exemplified contrastively. Sources have
been indicated except in the event of a few independent translations, which have been based on analogy with other examples.

As mentioned above, if no explicit or implicit contextual information is provided, both languages allow the past and the present perfect. The past is used when the language user is concerned with THEN, whereas the present perfect is used when the language user is concerned with NOW.

\[(1.21)\]
(a) *Kees turned off the radio and rolled a cigarette.* (Fahringer, 1999, p. 65)
(b) *Kees zette de radio af en rolde een sigaret.* (Fahringer, 1999, p. 65)

\[(1.22)\]
(a) *Have you seen my cat?* (Aarts & Wekker, 1993, p. 215)
(b) *Heb je mijn kat gezien?* (Aarts & Wekker, 1993, p. 215)

\[(1.23)\]
(a) *He has already read the book.* (Fahringer, 1999, p. 68)
(b) *Hij heeft het boek al gelezen.* (Fahringer, 1999, p. 68)

Once explicit or implicit contextual temporal information becomes retrievable, semantic alignment must be achieved. As far as semantic alignment with past-zone adverbials is concerned, English does not allow the use of a present perfect. Dutch, on the other hand, often does allow the use of a present perfect in past-zone contexts.

\[(1.24)\]
(a) *The train left at two o’clock.* (Fahringer, 1999, p. 69)
(b) *De trein is om twee uur vertrokken.* (Fahringer, 1999, p. 69)

\[(1.25)\]
(a) *Jane frequently walked to work in those days.* (Aarts & Wekker, 1993, p. 212)
(b) *Jane liep in die tijd vaak naar haar werk./*Jane is in die tijd vaak naar haar werk gelopen. (Aarts & Wekker, 1993, pp. 212–213)

\[(1.26)\]
(a) *Vijft准入 jaar lang hebben de paleontologen gespeculeerd over de vraag welke collega de schedelfragmenten van een primitief mens (tussen 1912 en 1914 ontdekt in een steengroeve in Sussex) heeft vervld.* (van Brederode & Koopman, 1990, p. 128)
(b) *For twenty-five years paleontologists have speculated about the question of which colleague faked the skull fragments of a primitive human being, discovered in a quarry in Sussex between 1912 and 1914.* (van Brederode & Koopman, 1990, p. 128)

\[(1.27)\]
(a) *The Romans founded York.* (van Brederode & Koopman, 1990, p. 128)
(b) *De Romeinen hebben York gesticht.* (van Brederode & Koopman, 1990, p. 128)
If multi-zone adverbials are used with a past, they will automatically receive a past reading. Once again, English will use a past to establish a past-zone context whereas Dutch may use either a past or a present perfect to establish the past-zone context.

(1.30) (a) He cycled to school every day. (Mackenzie, 1997, p. 40)
    (b) Hij fietste iedere dag naar school. / Hij is iedere dag naar school gefietst.
        (Mackenzie, 1997, p. 40)

(1.32) (a) We had a very busy time in the office this week. (a remark made on a Friday)
        (Ungerer, 2000, p. 151)
    (b) We hebben het heel druk gehad op kantoor deze week. / We hadden het heel druk op kantoor deze week.

1.5 Conclusion

In this first chapter, I have contextualized the grammatical category of tense from a mainly linguistic point of view. The overall goal was to provide a linguistic foundation for investigating the target structures of interest in this doctoral dissertation: the past and the present perfect when used to locate bygone situations in English and in Dutch. I started with a number of existing definitions of tense to focus on recurring items of both agreement and disagreement. The selected definitions highlighted two items for which there appears to be consensus: (1) tense is generally defined as a grammatical category of the verb and (2) there is a relationship between tense and time, which appears to be fluid and bidirectional, meaning that one tense may express several temporal meanings and one temporal meaning may be expressed by several tenses. However, the definitions also referred to items of disagreement, which are often addressed in discussions of the grammatical category of tense. The first item of disagreement which was investigated was a purely formal one and was related to the morpho-
logical nature of certain verb forms. Synthetic verb forms (e.g., present, past) appear to be more easily accepted as tense forms than analytic verb forms (e.g., will-future, present perfect). The second item of disagreement relates to the aspectual features of specific verb forms (e.g., will-future, present perfect) which would disqualify them as candidates for tense forms in their own right. Both items were discussed and it was decided to follow the characteristics of Declerck’s (1991, 2003, 2006) descriptive theory of tense in English, which accepts that certain verb forms (e.g., present perfect), even though they are analytic verb forms, are tenses in their own right since they have temporal structures which are unique and cannot be found in any other verb form in the English tense system.

Subsequently, I tackled the challenge of conceptualizing tense. Since tense is such a broad grammatical category, conceptualizing temporal features of the verb phrase in English may be a highly complex endeavour at times. Declerck’s descriptive theory of tense in English was used to first introduce the basic features of tense (e.g., linguistic time, temporal zero point, time-spheres, time zones). This was followed by a more detailed discussion of the specific features involved in conceptualizing the past and the present perfect in English to locate bygone situations (e.g., actualization focus, semantic alignment, temporal adverbials). The overall conclusion was that both the past and the present perfect may be used in English to locate bygone situations. In clauses without temporal adverbials, the past is used when the speaker is concerned with THEN, the present perfect is used when the speaker is concerned with NOW. In clauses with temporal adverbials, the speaker must ensure that semantic alignment is achieved between the tense that is used and the temporal adverbials, which provide crucial information regarding the contextual temporal information.

The last section in this first chapter was dedicated to the past and present perfect in Dutch. Both tenses in Dutch show many form-related, meaning-related and use-related similarities with the past and the present perfect in English. There are, however, also some fundamental differences. One of the most striking differences, for example, may be found in the possible co-occurrence of tenses and temporal adverbials. Whereas English does not allow the use of past-zone adverbials with a present perfect (e.g., *I have seen her yesterday versus I saw her yesterday), Dutch often does and even prefers the present perfect to the past on many occasions (e.g., Ik heb haar gisteren gezien/Ik zag haar gisteren).

The next chapter will continue the discussion of the FMU mappings under investigation in this doctoral dissertation but will move the focus to aspects of temporal complexity. The choice between a past and a present perfect to locate bygone situations in English is indeed a complex one, which many ESL learners—including Dutch-speaking ESL learners—grapple with when learning English. It is a choice which is generally highlighted in grammars of English. Many grammars of English for foreigners discuss the choice between the past and the present perfect in English. However, the complexity-inducing factor is
generally identified as interlingual differences between the language learners’ native languages and English. But is negative crosslinguistic influence the only factor that contributes to (SLA) complexity? Chapter 2 will shed light on the complex choice that L2 learners are forced to make and will provide a more comprehensive account of both complexity-related SLA issues at play in this choice and the contributors to complexity.
CHAPTER 2
TENSE AND COMPLEXITY

Fools ignore complexity. Pragmatists suffer it. Some can avoid it. Geniuses remove it.
(Alan Perlis)

2.1 Introduction
In her book *Biting the Wax Tailpole: Misadventures of an Armchair Linguist*, Elizabeth Little discusses how various languages in the world apply an abundance of diverging linguistic resources to allow language users to communicate—presumably accurately, meaningfully and appropriately—with other language users. At the end of the discussion of the linguistic resources used in relation to nouns, Little (2007) draws the following general conclusion about (language) complexity: “In language, as in life, complexity can be overwhelming at first. But under closer inspection, it can prove to be something sublime” (p. 54). Although Little makes no explicit academic references to the quantitative and qualitative features of the concept of complexity, her readers are able to retrieve the general gist of her argument from both the context in which the concept is used and—most likely—from their own experiences with (foreign) languages. Complexity is a feature with which both instructed and uninstructed L2 learners are often confronted when faced with the challenge of acquiring specific L2 target features such as aspect, case, modality, mood and tense, to name but a few.27 However, it is also a feature about which many L2 learners have their own (subjective) beliefs.

The aim of this chapter is to shed light on the concept of complexity by discussing (mainly) qualitative features of tense complexity in the context of second language acquisition (SLA). After all, it is the complexity that Dutch-speaking ESL learners experience with temporal FMU mappings in the English verb phrase which led to the selection of the L2 target features under investigation in this doctoral dissertation. Consequently, a better understanding of complexity is imperative for a thorough understanding of the research approach adopted in the studies reported on in Chapters 4, 5 and 6. Many lay people have preconceived ideas about complexity and often reduce tense complexity for (Dutch-speaking) ESL learners to overt instances of negative L1

---

27 In the discussion of complexity in this chapter, the terms *ESL learner* and *Dutch-speaking ESL learner* will be used interchangeably since many of the statements that are made in this chapter apply to ESL learners of varying L1 backgrounds and are not exclusively experienced by Dutch-speaking ESL learners. When placing the focus on the participants in the experiments, I will make explicit references using the term *Dutch-speaking ESL learners* if confusion may arise.
transfer, with ESL learners erroneously applying L1 processing strategies and/or L1 features to their English interlanguages, resulting in ungrammatical temporal L2 constructions (e.g., *I have spoken to her two weeks ago (as opposed to the grammatical option I spoke to her two weeks ago), which—according to many lay people—is purely the result of the influence of the Dutch sentence Ik heb twee weken geleden met haar gesproken).

The term which will be used to refer to tense complexity as experienced by L2 learners is temporal SLA verb-phrase complexity. In the first section of this chapter (Section 2.2), an operational definition of temporal SLA verb-phrase complexity will be provided, bearing in mind the tripartite approach to grammar (form, meaning and use), which was highlighted in the introductory chapter. The second section (Section 2.3) will consist of more judicious discussions of various terminological and conceptual considerations which have been taken into account for the operationalization of complexity in this doctoral dissertation. In this section, it will become apparent that the operational definition of complexity in the first section is motivated and influenced by findings from various linguistic subfields (e.g., linguistics, psycholinguistics, second language acquisition). The third and final section (Section 2.4) will link up the definition with the acquisition and instruction of the L2 target features under investigation: the past and present perfect when used to locate bygone situations in present-day English. In other words, the third section will provide an overview of the discussion of the temporal verb-phrase complexity that Dutch-speaking ESL learners may experience when trying to acquire the selected L2 target features under investigation in instructional settings.

This chapter should by no means be viewed as an exhaustive discussion of (temporal) SLA verb-phrase complexity in present-day English. Rather, it is meant as a nuanced discussion which will facilitate the identification of features that play a role in the acquisition and instruction of complex temporal verb-phrase morphology in present-day English in instructionally explicit settings. Insights into the broader concept of SLA complexity in general and into the more specific concept of temporal SLA verb-phrase complexity in particular will prove to be beneficial to a variety of methodological and pedagogical decisions taken with regard to the acquisition and instruction of temporal verb-phrase morphology in this dissertation.

2.2 Operationalizing temporal SLA verb-phrase complexity

The idea of terminological polysemy plays an important role when operationalizing concepts. The concept of complexity in the linguistic subdomain of SLA is no exception in this respect. Haspelmath (2009) describes the phenomenon of terminological polysemy, and although he relates it to the linguistic study of the grammatical category of case, he does admit that the concept of terminological polysemy may be found “in all areas of grammar” (p. 505):
As in all areas of grammar, the terminology surrounding case phenomena is often not straightforward: Linguists with different backgrounds use the same terms for somewhat or radically different concepts, or they use different terms for very similar or identical concepts. It is unlikely that terminological consensus will emerge soon, primarily because there is no consensus about the concepts that we need, and terminological polysemy will continue to be rampant because there are many more concepts than handy terms. But it is useful to be aware of some of the most important terminological issues. (p. 505)

The situation surrounding the concept of complexity in the specialized subdomain of SLA bears many resemblances to the terminological situation surrounding the grammatical category of case described by Haspelmath above. The aim of this first section is to make available an explicit and focused operational definition of temporal SLA verb-phrase complexity, which will be used to develop the remaining sections in this chapter.

For the discussion of what will be termed relative, temporal SLA verb-phrase complexity in English in this doctoral dissertation, the following definition may be drawn up:

Temporal SLA verb-phrase complexity is defined as the degree of local, objective and psycholinguistic verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, objective and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

Note that the definition provided above is a definition of relative, temporal SLA verb-phrase complexity in present-day English. Other scholars will undoubtedly operationalize the concept differently depending on a variety of factors such as the level of L2 proficiency, the research approach, the research design and methodology, and the selected L2 target features. The definition provided will be explained in detail in the remainder of this chapter and will reappear several times with features highlighted as the discussion of complexity in this chapter progresses.

In the definition above, the following general definitional features may be recognized:

1. Complexity is defined as a concept of degree and not as a purely categorical concept which reflects the opposition complex versus non-complex. Consequently, complexity may be placed on a complexity
continuum with certain L2 target features displaying a higher degree of complexity than others.28

2. Complexity is defined as a multifaceted concept. The various features of the concept that have been selected (e.g., tripartite approach to grammar, Dutch/English language pairing, absolute versus relative complexity, global versus local complexity) were chosen based on considerations made in this doctoral dissertation. These definitional features will be explained in the remainder of this chapter.

3. The choice of the L2 target features (i.e., complex instantiation of tense in present-day English) has been integrated into the operational definition bearing in mind both the theoretical tense-related issues explained in Chapter 1 and Ellis’s (2006, 2008) references to the feature of problematicity as the key criterion in selecting L2 target features in many SLA studies investigating the effects of instruction.29

The remainder of this chapter will investigate in more detail the qualitative features of relative, temporal SLA verb-phrase complexity. In the discussion of complexity, the following two main aspects will be featured: (1) aspects of linguistic complexity and (2) aspects of SLA complexity. Although most of these features will be discussed separately, they do interact with each other on a variety of levels.

2.3 Complexity: A multifaceted concept

2.3.1 Linguistic complexity

The concept of linguistic complexity is one which has been around for a long time. However, it has proven to be an extremely elusive idea to conceptualize, operationalize and discuss both objectively and unproblematically. This section will look at some of the contemporary attempts which have been made in the field of linguistics to discuss—mainly qualitatively—linguistic complexity. In addition to a discussion of the features often highlighted by linguists to approach the topic of linguistic complexity, an overview will be provided, which will serve as a concise summary of the main controversies that the discussion of linguistic complexity has generally centred on in the more general field of linguistics.

---

28 The idea of a complexity continuum brings with it the insurmountable difficulty of establishing metrics to ‘calculate’ quantitative features of complexity in an attempt to ascertain accurate degrees of complexity. Because of the problems related to establishing exact specifications about quantitative features of (language) complexity (see, for example, Brown, 2000, DeKeyser, 2005, Robinson, 1996), the focus in this chapter will be on qualitative features of complexity, which are relatively more accessible than quantitative features.

29 Grammatical complexity is one of several factors available to determine problematicity and, consequently, to select specific L2 target features for investigation in SLA studies. Others include, for example, acquisition sequences, linguistic theory and psycholinguistic theory (Ellis, 2008).
Over the last five years a renewed interest in complexity has become clearly visible in the linguistic landscape (e.g., Dahl, 2004, 2007, 2008; McWhorter, 2007, 2008a, 2008b; Miestamo, 2006, 2008; Miestamo, Sinnemäki, & Karlsson, 2008; Sampson, Gil, & Trudgill, 2009). Although the actual concept of linguistic complexity is by no means a new topic of discussion, for decades many linguists appeared to shy away from referring to linguistic complexity in most (cross)linguistic investigations. In so doing, they avoided any references to possible views which even remotely expressed that there were languages which were more complex—and, hence, linguistically more superior—than others. Such views were overtly expressed in the nineteenth century in nationalistic/Eurocentric assumptions (Kusters, 2008).

For a long time, many linguists appeared adamant in stressing a so-called principle of equi-complexity (e.g., Aitchison, 1981; Crystal, 1997), which states that the appearance of linguistic complexity in one domain of a language is counterbalanced by the disappearance of linguistic complexity in another domain of that same language. Such a form of complexity-based interaction leads to a situation in which languages see only little to no difference in (overall) linguistic complexity and remain—theoretically speaking at least—at the same level of linguistic complexity (Kusters, 2008). The principle of equi-complexity is extremely appealing in that it allows for a tremendous amount of scope (both intralinguistically and interlinguistically) when looking at linguistic complexity and the areas affected by the appearance and disappearance of linguistic complexity. In other words, in theory, the principle can be applied to a variety of structures—individually or additively—within one language and across a selection of languages. At the same time, however, the principle of equi-complexity does indeed pose a fundamental challenge to anyone on the lookout for evidence against the principle: How can one falsify such an all-inclusive principle (Kusters, 2008)?

Contemporary (cross)linguistic discussions of complexity are no longer hesitant in dealing with linguistic complexity and approach the subject head-on, providing, at times, ample resistance to the principle of equi-complexity referred to above (e.g., Dahl, 2004, 2008; Kusters, 2003, 2008; McWhorter, 2007, 2008; Miestamo, Sinnemäki, & Karlsson, 2008; Sampson, Gil, & Trudgill, 2009). What most of these discussions have in common is that they by no means claim that certain languages are complex and others are not. The general assumption in these discussions is that all languages show some form of complexity but that the levels of complexity need not be the same in all languages, resulting in languages or specific language features which are inherently more complex than others. The move from avoiding linguistic complexity to embracing and developing the concept of linguistic complexity in all of its attested

---

30 The principle of equi-complexity is also known as the principle of invariance of language complexity or the principle of complexity invariance (Sampson, 2009).
variety has resulted in both fruitful research and interesting findings, which are useful for research design and methodology drawn up to investigate complexity. The remainder of this section on linguistic complexity will investigate two distinctions which are often made in discussions of linguistic complexity: (1) absolute versus relative linguistic complexity and (2) global versus local linguistic complexity.

### 2.3.1.1 Absolute versus relative linguistic complexity

One of the first oppositional distinctions that can be found in the modern (crosslinguistic) literature on complexity is the distinction between absolute and relative linguistic complexity (see, for example, Kusters, 2008, Miestamo, 2006, 2008), which was initially highlighted and coined by Matti Miestamo in his 2006 publication *On the Feasibility of Complexity Metrics*. Since its introduction, the oppositional dichotomy of absolute and relative complexity has been keenly used and discussed in the literature on linguistic complexity.

The fundamental difference between absolute and relative linguistic complexity lies in the fact that the concept of absolute linguistic complexity is used to refer to linguistic complexity that is “not related to the experiences of a particular kind of language user” (Kusters, 2008, p. 4, quoted with original highlighting) but rather to linguistic complexity that is considered to be “an aspect of a language as an autonomous entity” (Kusters, 2008, pp. 4–5). In effect, this means that absolute linguistic complexity can be referred to as system-oriented or theory-oriented linguistic complexity. By extension, relative linguistic complexity can be defined as linguistic complexity relative to someone or something. Consequently, relative linguistic complexity can be referred to as user-oriented linguistic complexity. Both absolutist and relativist approaches to linguistic complexity offer a myriad of possibilities for treating and discussing complexity in language(s). However, the fundamental difference between the two positions is that absolutist approaches to linguistic complexity investigate linguistic complexity using a conceptual framework (i.e., a description and/or theory) of language that is said to be independent of any specific language user and/or language aspect whereas relativist definitions of linguistic complexity do not.

The seemingly unproblematic distinction between absolute and relative linguistic complexity is, however, not as straightforward as it appears. Kusters (2008) highlights some fundamental problems with it, one of which being the premise of the absolutist approach to linguistic complexity: its so-called unrelatedness to language users. Kusters (2008) is critical of the purported unrelatedness and he views the absolutist position as “a relativist position in disguise” (p. 8) since it is relative depending on the linguistic theory which is used to evaluate linguistic complexity. The problem related to operationalizing and discussing the concept of absolute linguistic complexity becomes even more
acute when it is acknowledged that a linguistic theory is, in fact, a model of a language user, albeit a possibly idealized language user (Kusters, 2008).

Even though the distinction between absolute and relative (linguistic) complexity is commonly found in the modern literature on linguistic complexity, it is not a distinction to which all linguistic scholars adhere. Dahl (2004, 2007, 2008) clearly expresses his reluctance to use the term relative complexity and prefers the terms cost, difficulty and demandingness to refer to what he calls “different aspects of ‘complexity for a user’” (2007, p. 42) (i.e., relative linguistic complexity in the current discussion of linguistic complexity). In Dahl’s terminology, complexity inevitably refers to absolute complexity and not to relative complexity. Although Dahl explains his reasons for keeping absolute linguistic complexity and relative linguistic complexity clearly separated, he provides somewhat inconsistent and contradictory guidelines as to how complexity-related terminology should be used. On the one hand, he expresses his reluctance to use, for example, the term relative (linguistic) complexity but, on the other hand, he uses the term agent-related complexity (i.e., relative linguistic complexity) when, in fact, he is referring to the three terms cost, difficulty and demandingness referred to above (Dahl, 2007). In addition, Dahl (2007) writes about keeping absolute and agent-related complexity apart and not identifying the (absolute) complexity of a language with difficulty. However, surely such an oppositional distinction is possible—even when using the term relative complexity—as long as clearly demarcated definitions are used whenever linguistically viable.

Taking into account the controversy surrounding the terms absolute (linguistic) complexity and relative (linguistic) complexity, the term relative complexity will be used in the discussion of both linguistic complexity and SLA complexity in this doctoral dissertation since it provides a succinct and linguistically viable way of referring to both linguistic complexity and SLA complexity, which are major determinants of the complexity in the research at hand: temporal verb-phrase complexity in present-day English relative to Dutch-speaking ESL learners in instructionally explicit settings. This is reflected in the definition of complexity as follows:

Temporal SLA verb-phrase complexity is defined as the degree of local, objective and psycholinguistic verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, objective and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

31 For more elaborate definitions and discussions of the terms cost, difficulty and demandingness see Dahl, 2004, 2007.
2.3.1.2 Global versus local linguistic complexity

A second distinction that is made in the linguistic discussion of complexity is the distinction between global linguistic complexity and local linguistic complexity. The concept of global linguistic complexity is used when talking about overall linguistic complexity in a language, in other words, linguistic complexity across the entirety of a language system. Local linguistic complexity, on the other hand, focuses on the linguistic complexity of a part (or subsystem) of the entire system. Examples of local complexity are case complexity, mood complexity and temporal complexity, which all focus on subsystems of a larger system. Although many linguistic scholars have avoided discussing overall linguistic complexity in the past, and have, instead, preferred to adhere to the principle of equi-complexity (see Section 2.3.1), there is currently a growing interest in both global complexity and the development of accessible and feasible complexity metrics to map (features of) global linguistic complexity both quantitatively and qualitatively (see, for example, McWhorter, 2007, 2008b). In theory, such metrics would allow crosslinguistic comparisons of global complexity.

At the same time, however, there are scholars who regard any attempt to study global linguistic complexity as problematic. In their discussion of global linguistic complexity, they highlight that researchers need to take into account some fundamental problems when trying to operationalize and discuss global linguistic complexity. Miestamo (2008) discusses two such problems, which he refers to as the problems of representativity and comparability. According to Miestamo (2008), these two problems loom in any study of global complexity and he summarizes them as follows:

The problem of representativity means that no [complexity] metric can pay attention to all aspects of grammar that are relevant for measuring global complexity. Even if this were theoretically possible, it would be beyond the capacities of the mortal linguist to exhaustively count all grammatical details of the languages studied, especially in a large-scale cross-linguistic study. (p. 30)

The problem of comparability is about the difficulty of comparing different aspects of grammar in a meaningful way, and especially about the impossibility of quantifying their contributions to global complexity. (p. 30)

The problem of comparability is also the main reason why Miestamo (2008) clearly states that “the cross-linguistic study of grammatical complexity should primarily focus on specific areas of grammar, i.e., on local complexity” (p. 31). In other words, Miestamo clearly advocates the study of local linguistic complexity. In this respect, the study of relative, temporal SLA verb-phrase complexity in present-day English is in line with Miestamo’s (and other scholars’) views and provides ample focus on two complex instantiations of tense in English. As such, the discussion at hand clearly focuses on local complexity. The L2 target features under investigation are problematic instantiations of the gram-
matical category of tense in English, more specifically the use of the past and the present perfect when used to locate bygone situations in present-day English. The grammatical category of tense is generally expressed in English in the verb phrase but often shows forms of interplay with other temporal features in discourse (e.g., adverbials). As such, temporal SLA verb-phrase complexity in English is an example of local complexity since it is found only in the verb phrase. However, other parts of the grammar of English (e.g., adverbials) clearly interact with verb-phrase elements. Both of these aspects (i.e., tense and context) will have to be considered in a judicious discussion of relative, temporal SLA verb-phrase complexity in English. The feature of local linguistic complexity is incorporated into the definition of complexity as follows:

Temporal SLA verb-phrase complexity is defined as the degree of local, objective and psycholinguistic verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, objective and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

2.3.2 Complexity in SLA research

So far, the discussion of complexity has focused on linguistic complexity, but the field of SLA has also turned to complexity in various lines of research. Over the last decade and a half a growing interest in complexity in SLA research has made itself visible. This interest has been attested in the increasing number of SLA-oriented publications dealing—partially or exclusively—with the concept of complexity (e.g., de Graaff, 1997; de Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2007; De Visscher, 2005; DeKeyser, 2005; Ellis, 2008; Ellis & Barkhuizen, 2005; Housen, Pierrard, & Van Daele, 2005; Hulstijn & de Graaff, 1994; Kempe & Brooks, 2008; Larsen-Freeman, 1997, 2003; Larsen-Freeman & Cameron, 2008; Skehan & Foster, 2007; Spada & Tomita, 2007). The increasingly popular use of the concept of complexity in the linguistic subdomain of SLA has led to a number of domain-specific conceptual and operational definitions of complexity. This is why the decision was taken to discuss the concepts of linguistic complexity and complexity in SLA settings separately in this doctoral dissertation. Of course, as already mentioned, a distinct demarcation of the borders between linguistic complexity and SLA complexity is untenable since interaction and consequent overlaps are inevitable.

2.3.2.1 Complexity as a dependent or independent variable

In their paper presented at the 2008 Anélia conference in Leiden, Housen, Van Daele and Bulté (2008) discuss the intricacies involved in investigating com-
plexity in an SLA context. Housen et al.’s (2008) paper makes clear that the concept of complexity in SLA research is multifaceted and offers many fruitful definitional and methodological possibilities. However, before discussing the (theoretically) demarcated forms of complexity, one research-related issue regarding the operationalization of complexity in SLA studies merits further discussion. In essence, the concept of complexity generally functions as a variable in SLA research and may do so in two conceptually and operationally distinct ways: (1) as a dependent variable (e.g., Derwing & Rossiter, 2003; Ellis, 2008; Ellis & Barkhuizen, 2005; Norris & Ortega, 2000; Spoelman & Verspoor, 2009) or (2) as an independent variable (e.g., DeKeyser, 2005; Ellis, 2008; Norris & Ortega, 2000; Spada & Tomita, 2007; Yuan & Ellis, 2003).

As an example of complexity as a dependent variable, Housen et al. (2008) state that complexity is generally operationalized as a basic dimension of L2 proficiency and L2 performance, often together with the concepts of accuracy and fluency. Both Richards and Schmidt (2002) and Thornbury (2006) provide definitions of complexity as a dependent variable:

A composite measure of language use, normally reflecting the length of utterances and the amount of subordination used. In studying a second language learner’s discourse or interlanguage complexity is one measure of L2 development. (Richards & Schmidt, 2002, p. 96)

A learner’s language is complex if it uses structures more typical of advanced learners than of lower level learners. A learner may be both accurate and fluent, but if their output consists of very simple sentences, they cannot really be said to be advanced. Factors that are taken into account when assessing complexity include:

- the amount of subordination, including the use of complex sentences
- the use of pronouns for back reference
- the proportion of lexical verbs to linking verbs: the more of the former, the more complex
- the proportion of content words to function words: the more of the former, the more complex
- the frequency of use of conjunctions. (Thornbury, 2006, p. 40)

Following this line of operationalization, the dependency of L2 complexity on learner-internal and learner-external factors has proven to be a highly productive line of SLA research.34

32 The acronym Aneía stands for Association néerlandaise de linguistique appliquée, which is the official name of the Dutch association for applied linguistics. More detailed information may be found online at www.anela.nl.

33 For a more thorough discussion of general and experiment-specific variable features, see Chapters 4, 5 and 6.

34 Examples of learner-internal factors are age, aptitude, attitude, awareness, cognitive style, interest, learning strategies, motivation, personality, prior experience, sex, working memory. Examples of learner-external factors are feedback, input frequency, input quality, input quantity,
As an example of complexity as an independent variable, Housen et al. (2008) focus on instruction studies, which generally investigate the impact of the complexity of one or several L2 target features on the effects, effectiveness and efficiency of L2 instruction. The definition of complexity used in this doctoral dissertation operationalizes complexity as an independent variable. The choice of the L2 target structures under investigation was based on the high degree of problematicity that the target structures appear to be associated with in grammars of English. Its integration into the definition of complexity used in this doctoral dissertation is as follows:

Temporal SLA verb-phrase complexity is defined as the degree of local, objective and psycholinguistic verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, objective and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

However, caution should be exercised with the distinction between complexity as a dependent and an independent variable. The dichotomous operationalization of linguistic complexity as either a dependent or an independent variable in SLA research may create the impression that we are dealing with a clearly defined oppositional phenomenon. In reality, however, this opposition is not always as clear-cut as it initially appears. If a specific (grammatical) L2 target feature is categorized as a complex L2 target feature (i.e., an independent variable) and is selected for, for example, instruction, one would logically presume that the goal of instruction is the acquisition of that complex L2 target feature. However, the goal of instruction of a complex L2 target feature would be to bring about consequent changes in the L2 learners’ interlanguages, which, in turn, would be characterized by a higher degree of interlanguage complexity (i.e., a dependent variable) as a result of acquiring a complex L2 target feature. Such logic would lead to circular reasoning and would defy any concrete form of operationalization of complexity. To avoid such reasoning, the distinction between complexity as a dependent variable and an independent variable will be based in this doctoral dissertation on the initial focus of research interest. Since the initial focus of research interest is the selection of problematic L2 target features for acquisition and instruction, complexity is defined as an independent variable.

---

2.3.2.2 The absolute—relative linguistic complexity distinction applied in SLA research

As was already mentioned, the form of complexity that is the focus of investigation in this doctoral dissertation is what Miestamo (2006) calls relative complexity, that is, complexity relative to someone or something. In their discussion of complexity, Housen et al. (2008) refer to this form of complexity in SLA research using the terms cognitive complexity or difficulty, which—according to them—comprises two forms of difficulty: (1) objective and (2) subjective difficulty. Housen et al. (2008) define the term cognitive complexity as the ease or difficulty with which L2 learners process language items in acquisition and use. Although both forms of difficulty (objective and subjective) may provide substantial contributions to a better understanding of both L2 acquisition and L2 instruction, the focus in this doctoral dissertation is on objective difficulty simply because this form of difficulty allows for a greater degree of research control. Subjective difficulty is largely—though not exclusively—the result of learner-internal factors, which are often beyond researchers’ direct control and extremely difficult to distil from L2 learners’ performance in classroom or laboratory settings. Furthermore, Housen, et al. (2008) provide a non-exhaustive selection of objective determinants (i.e., learner-external factors), which are factors that objectively determine the cognitive difficulty of specific L2 target features. Some of their more prominently featured objective determinants are input saliency, pedagogical rule complexity, processing costs and linguistic complexity.

The distinction between absolute and relative complexity was discussed above and although it may have been presented as an either-or distinction, in practice, some discussions of complexity appear to resist—partially or completely—such a discrete distinction. The discussion of relative, temporal SLA verb-phrase complexity in present-day English could be interpreted as providing resistance to such a clear distinction in that certain issues of temporal SLA verb-phrase complexity in English (e.g., L1-induced complexity) can be said to be clearly relative, that is, specific to a certain ESL learner or to a specific group of ESL learners. Other issues of temporal SLA verb-phrase complexity in English (e.g., general problems related to the demarcation of semantic boundaries in the English tense system) can be said to be more universal or absolute, in that they are experienced by ESL learners in general and are related to inherent features of the language system. However, for the discussion at hand the absolute—relative distinction will be used to allow for a clearer discussion of the complexities at hand. Certain issues related to relative, temporal SLA verb-phrase complexity are definitely universal but interaction with more relative

---

35 Dahl’s (2007) reluctance to use the term complexity to refer to relative complexity does not appear to be of any direct concern to Housen et al. (2008), who appear to use the terms (cognitive) complexity and difficulty interchangeably.
issues will also be highlighted in the discussion below. The idea of objective complexity is incorporated into the definition of complexity in this doctoral dissertation as follows:

Temporal SLA verb-phrase complexity is defined as the degree of local, **objective** and psycholinguistic verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, **objective** and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

### 2.3.2.3 Complexity and the acquisition of L2 mappings

Complexity as a dependent or an independent variable has been the focus of discussion in the subdomain of SLA for some time, where it is generally referred to as complexity or difficulty. The idea of difficulty is discussed, for example, by Robert DeKeyser in his 2005 publication *What Makes Learning Second-Language Grammar Difficult: A Review of Issues*. In his article, DeKeyser discusses how tricky the concept of difficulty is. He finally settles on three factors which—according to him—are “involved in determining grammatical difficulty” (2005, p. 3): (1) complexity of form, (2) complexity of meaning and (3) complexity of the form-meaning mapping. DeKeyser (2005), however, is extremely careful in defining difficulty. He states that a ternary division of grammatical difficulty is not complete in that it omits any references to “the core psycholinguistic difficulty of acquisition, that is, the difficulty of grasping the form-meaning relationship while processing a sentence in the L2” (2005, p. 3), which—according to DeKeyser—is determined by the transparency of the form-meaning relationship to L2 learners processing language for meaning (at least for L2 learners who do not receive any explicit instruction on the rules governing the form-meaning mappings). DeKeyser’s statement about psycholinguistic difficulty is a valid one, which will be taken up when the concept of complexity is discussed with reference to the temporal L2 target features under investigation in this doctoral dissertation (see Section 2.4). For the time being, the idea of psycholinguistic complexity will be taken up in the definition of complexity as follows:

Temporal SLA verb-phrase complexity is defined as the degree of local, **objective** and **psycholinguistic** verb-phrase problematicity experienced by Dutch-speaking adult ESL learners who are receiving or have received explicit instruction aimed at acquiring temporal verb-phrase FMU mappings in present-day English. The concept of local, **objective** and **psycholinguistic** verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

---

36 DeKeyser (2005) appears to use the terms *complexity* and *difficulty* either interchangeably or indiscriminately. However, the reader is advised to bear in mind that the terms *complexity* and *difficulty* that DeKeyser refers to in his 2005 publication are, by and large, synonymous with the term *objective cognitive difficulty* referred to by Housen et al. (2008).
objective and psycholinguistic temporal verb-phrase problematicity is defined as the sum of form-related, meaning-related, use-related, mapping-related and other complexity-inducing-related factors experienced by the Dutch-speaking adult ESL learners referred to above.

DeKeyser (2005) does not appear to make any regular, explicit references to the grammatical dimension of use in his discussion of mapping difficulty but from his discussion of the various aspects of difficulty it is clear that the dimension of use is discussed together with the dimension of meaning, which is often the case in discussions of FMU mappings in the (SLA) literature. The distinction between meaning and use is not always easily discernable in the tense-aspect system (Celce-Murcia & Larsen-Freeman, 1999). Consequently, specific issues dealing with meaning-related and use-related complexity will often be discussed together. This is clearly visible in Section 2.4, which is completely devoted to the complex temporal FMU mappings under investigation.

DeKeyser (2005) appears to use a discrete separation of grammatical difficulty and “the core psycholinguistic difficulty of acquisition” (p. 3). However, such discrete separation of two closely related concepts is—in my opinion—viable only if one takes into account that certain areas of overlap may exist. Specific issues of complexity related to these areas of overlap should subsequently be discussed bearing in mind the actual overlaps that are present (see also Section 2.3.2). A concrete example may be required at this point to highlight the importance of such overlaps so let us have a look at formal regularity and irregularity with regard to the past in present-day English.

The regular past is formed in present-day English by adding the inflectional suffix -(e)d to the (present) infinitive form (i.e., the base or stem) of the verb. This process produces past forms such as analyzed [analyse + -ed], loved [love + -ed], texted [text + -ed], walked [walk + -ed]. In addition to regular past forms, present-day English also has verbs which have irregular past forms, which are formed using morphological processes such as ablauting (e.g., drove, ran), consonant alternation (e.g., lent, sent) and suppletive forms (e.g., was/were, went). The dichotomy of formally regular and irregular past verb forms may appear to be straightforward but it is a dichotomy which is often regarded as problematic since ESL learners not only need to acquire the individual forms of such verbs but also need to acquire which verb forms carry which meanings and how those forms and meanings are intertwined and put to use in present-day English.\(^\text{37}\) In addition, the regular verbs are not always as regular as one might expect. It is true that the majority of regular verbs form their past verb forms fairly mechanically by adding -(e)d to the (present) infinitive form. However,\(^\text{37}\)

\(^{37}\) Standard present-day English also has verbs which have both regular and irregular past verb forms. Such past verb forms may be conditioned, for example, geographically (e.g., drove (drew (BrE)/dove (AmE)), smelt (smelt, smelt (BrE)/smelled (AmE)), spell (spelt, spelled (BrE)/spelled (AmE)) or semantically (e.g., hung (hung/hanged), knit (knit/knitted), ring (rang/ringed)).
many regular verbs do in fact show additional formal changes such as final consonant doubling (begged [beg + g + -ed], occurred [occur + r + -ed] and final -y replacement (carried [carri + -ed], levy [levi + -ed], marry [marri + -ed]), which are not always characterized as straightforward changes (see Section 2.4.3 for further details). The conditioned spelling changes described above clearly contribute to what DeKeyser refers to as complexity of form. However, such specific spelling changes also contribute to the psycholinguistic difficulty of acquisition since they affect—to a certain degree—the predictability and ease or difficulty of FMU mappings. Consequently, the factor of psycholinguistic difficulty has also been integrated into the definition of complexity provided in Section 2.2.2.

DeKeyser continues his discussion of complexity/difficulty by highlighting form-related, meaning-related and mapping-related problems. When arriving at mapping complexity/difficulty, DeKeyser (2005) introduces the concept of transparency to refer to the link between form and meaning. He lists three factors which may cause a lack of transparency: (1) redundancy, (2) optionality and (3) opacity. He admits, however, that there may be more factors at play. To understand these three factors better, they will be highlighted in more detail below.

2.4 Relative, objective and psycholinguistic temporal SLA complexity and the English verb phrase

2.4.1 Introduction
The aim of this section is to focus on the actual temporal verb-phrase complexity that Dutch-speaking ESL learners grapple with when confronted with complex (problematic) temporal verb-phrase mappings in present-day English in instructionally explicit settings. In essence, we are dealing with a mapping, more precisely, a ternary mapping according to the adopted tripartite approach to grammar (Celce-Murcia & Larsen-Freeman, 1999; Larsen-Freeman, 1995, 2001, 2003). Following an adapted version of DeKeyser’s (2005) discussion of what makes learning L2 grammar difficult, at least five possible sources of relative, temporal SLA verb-phrase complexity in English may be identified: (1) form-related complexity, (2) meaning-related complexity, (3) use-related complexity, (4) mapping-related complexity and (5) other complexity-inducing factors (e.g., L1-induced complexity, outcome measure complexity). However, DeKeyser’s (2005) concern about the completeness of this division must also be addressed. Since the focus in this chapter is on relative, objective and psycholinguistic complexity in an SLA context, we must also take into account DeKeyser’s (2005) reference to the “core psycholinguistic difficulty of acquisition, that is, the difficulty of grasping the form-meaning relationship while processing a sentence in the L2” (p. 3) and factors which may influence the process of grasping mappings such as those suggested by DeKeyser (2005) (redundancy, optionality, opacity) and Housen et al. (2008) (input saliency, processing costs, linguistic complexity). I will assume in this doctoral dissertation that the degree
of psycholinguistic difficulty of acquisition is the result of the sum of form-related complexity (Section 2.4.3), meaning-related and use-related complexity (Section 2.4.4), mapping-related complexity (Section 2.4.5) and other complexity-inducing factors (Section 2.4.6), which all interact in a multitude of ways.

It should be reiterated that the various forms of complexity which will be discussed are not always discrete forms of complexity. Overlaps between the various forms are inevitably present. Consequently, this may lead to a possible blurring of established, demarcated forms of complexity. Nevertheless, in an attempt to maintain a certain degree of overview, every effort will be made to discuss the various forms of complexity separately even though this may not necessarily be an accurate reflection of the online processing challenges that ESL learners face when acquiring complex FMU mappings in instructionally explicit settings.

2.4.2 The complex L2 target features under investigation

Before embarking on a more detailed discussion of the various forms of complexity, let us look at the complex target feature under investigation to remind ourselves of the actual realizations of the complex L2 target features which are being investigated.

In the introduction to his publication *Meaning and the English Verb*, Geoffrey Leech comments on possible problems that ESL learners may face. He starts the introduction of the 2004 edition of his publication by stating the following:

> Every language has its peculiar problems of meaning for the foreign learner. Many people would agree that in the English language, some of the most troublesome yet fascinating problems are concentrated in the area of the finite verb phrase, including, in particular, tense, aspect, mood and modality. The goal of this book is to describe these fields of usage systematically and in some detail for teachers and advanced students of English as a foreign or second language. (p. 1)

Leech is not alone in suggesting that tense—in addition to a variety of other grammatical categories and linguistic phenomena—may cause (meaning-related) problems for ESL learners. In addition to finding such observations in many other publications of a more theoretical nature, they are also clearly present in more practical discussions of the English language in general and of English grammar in particular (e.g., Aarts & Wekker, 1993; Aitken, 1992; Burrough-Boenisch, 2004; Butlerman, 2007; Carter & McCarthy, 2006; Celce-Murcia & Larsen-Freeman, 1999; Cumps & Vekemans, 2005; Declerck, 1991, 2003, 2006; Dekeyser, Devriendt, Tops, & Geukens, 1999; Downing & Locke, 2006; Foley & Hall, 2003; Hannay & Mackenzie, 2002; Hoffmann & Hoffmann, 2001, 2005; Huddleston & Pullum, 2005; Mackenzie, 1997; Rijks, 2006, 2008; Ungerer, 2000; van Brederode & Koopman, 1990)

Although some of the cited references above are specifically aimed at Dutch-speaking ESL learners, many of the problems are more universal in that
they may occur for a variety of ESL learners. Consequently, they may be found in discussions aimed at both ESL learners in general and at specific groups of ESL learners who share a common feature (e.g., a common L1). It is the ubiquity of temporal problems for ESL learners which led to the choice of the L2 target features under investigation. In addition to presenting more universal (acquisitional) problems, the L2 target features under investigation are often said to pose a special learning challenge for Dutch-speaking ESL learners because of attested L1-induced problems with which many Dutch-speaking ESL learners of varying proficiency (i.e., beginner, intermediate, advanced) grapple.

The L2 target features under investigation are the English past and the English present perfect when used to locate bygone situations. Of special interest are the instantiations where a choice between these two tenses is required to locate bygone situations in past-zone contexts. In Declerck’s terminology (1991, 2003, 2006) a past-zone context is defined as a context which lies completely before the temporal zero-point ($t_0$) and is disconnected from $t_0$. As such, the L2 target features are only one part of the challenge that Dutch-speaking ESL learners face since the use of the English past to locate bygone situations often ‘competes’ with a second mapping, namely, the English present perfect, which may also be used to refer to bygone situations. In this respect, Leech’s (2004) choice of words is clear when he says that “it is worth making the point that the Present Perfect and Simple Past are not mutually exclusive choices: there are many situations where either of these tenses would be suitable” (p. 35). The idea of grammar as choice plays an important role in guiding (Dutch-speaking) ESL learners when they are faced with having to make a choice between the past and the present perfect in English.

The past and the present perfect may be semantic neighbours in the English tense system but they are clearly distinct semantically in that the past places the temporal focus on THEN whereas the present perfect places the temporal focus on NOW. This semantic difference often leads to restrictions on how the two tenses are used (e.g., with specific adverbials). The semantics of both tenses result in a choice between the past and the present perfect which is by no means arbitrary (see, for example, Ungerer et al., 2009). The lack of arbitrariness is reflected by Declerck (2006) when he says that the use of a specific tense “is wholly determined by its semantics (= temporal structure), which has to fit in with the temporal information given by the time-specifying adverbials or by the context” (p. 599). Additional information on the issues of meaning and use will be provided below in the discussion of meaning-related and use-related complexity. By way of example, consider the following sentences:

---

38 The idea of competition is one which features prominently in, for example, Bates and MacWhinney’s competition model (1989). As fruitful as it may be for the discussion at hand, Bates and MacWhinney’s idea of competition will not be pursued in this dissertation.
Both (2.1) and (2.2) may refer to the exact same situation, that is, the asking of a question, which took place at a specific point in time located before the temporal zero point. However, the two situations differ with respect to their temporal focus. The sentence containing the past (2.1) clearly puts the focus on THEN. Consequently, this temporal form is obligatory whenever the temporal focus is placed on THEN and when no connection with the present time-sphere can be retrieved. The sentence containing the present perfect (2.2) clearly puts the focus on NOW. This temporal form is obligatory whenever the temporal focus is placed on NOW, that is, when a connection with the present time-sphere can be retrieved.

If we were to use our ternary approach to grammar to represent these two tenses schematically, the temporal FMU mappings in (2.1) and (2.2) could be represented using Figures 2.2 and 2.3 respectively.
Figure 2.2. Form, meaning and use of the English present perfect tense when used to locate the bygone situation in the sentence *She has asked me a question*.

In keeping with the ternary approach to grammar used in this doctoral dissertation, the choice between the past and the present perfect may be viewed as a choice between two temporal FMU mappings. If we apply a ternary approach to the choice that ESL learners are faced with when locating bygone situations in present-day English, we notice that there are form-related, meaning-related and use-related overlaps and differences with regard to the temporal FMU mappings. However, the main problem that ESL learners face is relatively unambiguous when they are forced to make a choice between the past and the present perfect in past-zone contexts. According to the rules of standard present-day English, the present perfect is completely incompatible with past-zone contexts because of its focus on NOW. This rule can be found in most—if not all—discussions of the English tense system aimed at ESL learners in general (e.g., Aitken, 1992; Carter & McCarthy, 2006; Celce-Murcia & Larsen-Freeman, 1999; Declerck, 1991, 2003, 2006; Downing & Locke, 2006; Huddleston & Pullum, 2005) and at Dutch-speaking ESL learners in particular (e.g., Aarts & Wekker, 1993; De Moor, 1998; Koning & van der Voort, 1997; Mackenzie, 1997; van Brederode & Koopman, 1990).

The following sections will look at the complex temporal FMU mappings under investigation and will draw comparisons between the past and the present perfect. The present perfect is generally considered to be the main ‘rival’ or
‘competitor’ of the past when Dutch-speaking ESL learners are faced with having to choose the appropriate tense in, for example, past-zone contexts.

2.4.3 Form-related complexity
Using a ternary mapping approach (form, meaning, use), we can associate the dimension of form most easily with the concept of accuracy.\(^39\) One of the challenges that Dutch-speaking ESL learners face when dealing with English past and present-perfect verb forms is learning to produce formally correct (i.e., accurate) verb inflections. Ungrammatical past and present-perfect verb inflections such as *she beated, *we developped, *you have drank, *he panicked and *they runned are extremely common occurrences which are by no means restricted to instructionally explicit settings. As a result, formal tense-related issues are part of the discussion of the English tense system found in many grammars of English.

Of course, one could argue that such formally ungrammatical verb inflections are not always consistently detected and that detection depends—to a large extent—on whether such inflections are produced in the written mode or the oral mode (e.g., *we developped cannot be detected in the spoken mode). Nonetheless, producing formally grammatical verb inflections is a challenge which ESL learners face in both modes. Therefore, this problem merits further investigation.\(^40\) What is more, some of the formal challenges are closely linked to issues of meaning and use, with a selection of verb forms which are conditioned as a result of meaning-related issues (e.g., *ringed versus rang) or use-related issues (e.g., *dived versus dove). For the discussion at hand, the emphasis will be on the written mode since the participants in the experiments were all students who had enrolled in a three-year translation programme which would eventually lead to a bachelor’s degree in applied linguistics (with a specialization in translation). Consequently, the experiments and the instruction were designed with a clear focus on the written mode.

According to DeKeyser’s (2005) argument with regard to difficult L2 grammar, we could assume that any difficulty of form can be described as the “number of choices” (pp. 5–6) involved in selecting all the correct morphemes/allomorphs to express the meanings that one wishes to convey. DeKeyser (2005) continues his argument by stating the following:

\(^39\) In Larsen-Freeman’s (1995, 2001, 2003; Celce-Murcia & Larsen-Freeman, 1999) ternary conceptualization of grammar, which was introduced in the introduction to this dissertation, the dimension of form is associated with the concept of accuracy, the dimension of meaning with the concept of meaningfulness, and the dimension of use with the concept of appropriateness.

\(^40\) The fact that Dutch-speaking ESL learners face problems when producing formally grammatical verb inflections in both the written mode and the oral mode does not necessarily mean that these problems are identical. Written-mode problems (e.g., problems of spelling) and oral-mode problems (e.g., problems of pronunciation) may be totally different depending on the actual verb inflections which are produced. However, a clear distinction between both modes does not necessarily reflect the realities of the complex, multifaceted SLA process. ESL learners may be guided, for example, by issues of pronunciation when producing written verb inflections.
Clearly, this problem is most complex in richly inflected languages, whether they be agglutinative, polysynthetic, or inflectional in the narrow sense. Everything else (such as semantic difficulty) being the same, the more that needs to be expressed overtly, the more choices need to be made about morphemes, allomorphs, and their position. (p. 6)

I agree with the second part of DeKeyser’s statement (“Everything else (such as semantic difficulty) being the same, the more that needs to be expressed overtly, the more choices need to be made about morphemes, allomorphs, and their position”) but the fact that DeKeyser appears to equate rich inflection with a high, or at least higher, level of form-related complexity would appear to indicate that his discussion of difficulty of form is either synonymous with or heavily reliant on the quantity of choices. Nowhere in DeKeyser’s discussion of form-related complexity is there any (explicit) reference to the quality of the form-related choices that need to be made. A focus on quantity may lead to problems when trying to ascertain the nature and possibly the degree of formal complexity of temporal verb-phrase morphology in present-day English.

Pinker’s (1994, 1999) words speak volumes when he says that “the creative powers of English morphology are pathetic compared to what we find in other languages” (1994, p. 127) or “English inflection is famous among linguists for being so boring” (1999, p. 29), after which he makes a reference to the fact that English verbs come in only four forms: (1) the base form, (2) the -s form, (3) the -ing form and (4) the -ed form. A focus on quantity as far as inflectional tense morphology is concerned would definitely categorize present-day English as less complex than a whole range of other languages. Clearly, there is more at play than simply the quantity of choices that need to be made.

From a morphosyntactic point of view, the choice between the past and the present perfect (when used to locate bygone situations) in present-day English entails the choice between two tense forms which can be described formally by means of rules. The regular past in present-day English is generally formed by adding a past suffix (-ed) to the base form of the verb (e.g., allowed, greeted, played, warranted). In the case of irregular verbs, the resulting past verb inflections may display specific irregularities (e.g., ate, began, led, slit), which will be discussed in more detail below. The regular present perfect, on the other hand, is generally formed by combining the present of the auxiliary verb have with the regular past-participle form of the main verb (e.g., has/have allowed, has/have greeted).

---

41 It should be stressed that Pinker (1994, 1999) is referring to inflectional morphology in his statements about English and not to derivational morphology. His references are based on, for example, the number of inflectional verb forms available in English, which he compares with languages such as Greek, Italian, Spanish and Turkish, which are morphologically more prolific languages from an inflectional point of view.

42 The operationalization of various types of complexity (e.g., formal, functional) remains a problem in contemporary SLA research (see, for example, DeKeyser, 1998, Ellis, 1990, 1997, Krashen, 1981, for opposing views on the complexity of the third person simple present -s marker in English).
has / have played, has / have warranted). In addition to the formal irregularities which the auxiliary have displays in the present, the past participle of irregular verbs may increase the degree of overall irregularity of this tense (e.g., has / have begun, has / have eaten, has / have led, has / have slit). Once again, the degree of irregularity needs to be put into perspective. An increased degree of irregularity does not mean, for example, that a present perfect is inherently more complex to comprehend or produce than a past. I will come back to this idea later on when discussing the possible effects of frequency on the acquisition of irregular past and present-perfect verb inflections.

With regard to both past and past-participle verb inflections an important distinction is generally made in present-day English between regular and irregular verbs. The remainder of this section will first look at the formation of the past and the past-participle forms for regular verbs. Subsequently, the formation of the past and the past-participle forms for irregular verbs will be discussed. The formal discussion provided in this section is not meant to serve as an exhaustive overview of English verb inflections but rather as a backdrop for the general form-related challenges that ESL learners face. For more detailed information the reader is advised to consult publications dealing with English grammar in general or with tense in English in particular (e.g., Aarts & Wekker, 1993; Alexander, 1988; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Carter & McCarthy, 2006; Declerck, 1991, 2003, 2006; Greenbaum, 2000; Greenbaum & Quirk, 1990; Huddleston & Pullum, 2005; Koning & van der Voort, 1997; Ungerer, 2000).

The past and past-participle forms of regular verbs are identical and the form type used for these forms is the -ed form, which is one of the four morphological variants that regular verbs have in present-day English. The -ed form is generally constructed by adding the inflectional suffix -ed to the base form of the regular verb (e.g., allow / allowed, call / called, text / texted, watch / watched). Although the majority of regular verbs allow the computational addition of the inflectional suffix -ed to the base form of the regular verb, there are some extremely common spelling changes, which need to be taken into account when forming regular past and past-participle verb inflections. These spelling changes are relatively predictable. However, they nonetheless cause common formal

---

43 Although many irregular verbs display both irregular past verb inflections and irregular past-participle verb inflections (e.g., begin, eat, lead, ill), complete irregularity across both types of verb inflections is not a prerequisite for membership to the group of irregular verbs (e.g., sew / sewed (past)/sewed or sewn (past participle), show / showed / showed or shown, sow / sowed / sown or sown, strew / strewn / strewn or strown).

44 The distinction between regular and irregular verbs is also referred to using the terms weak verbs and strong verbs respectively (Crystal, 2003; Declerck, 2003, 2006).

45 The four morphological variants for regular verbs are the base form (e.g., allow, push, surf, travel), the -s form (e.g., allows, pushes, surfs, travels), the -ing form (used for the present participle form) (e.g., allowing, pushing, surfing, travel(l)ing) and the -ed form (used—in the case of regular verbs—for both the past and the past-participle forms (e.g., allowed, pushed, surfaced, travel(l)ed)).
problems for (Dutch-speaking) ESL learners. The spelling changes that occur in the formation of past and past-participle verb inflections are considered to be the results of general spelling rules in present-day English. As such, they are not unique to verb inflections and are also found in the formation of, for example, adjectives, adverbs and nouns. For the formation of the -ed form, the spelling changes may be categorized as follows: (1) final consonant doubling, (2) -e deletion and (3) final -y replacement.

**Final consonant doubling.** The first general spelling rule is the rule of final consonant doubling, which occurs only when specific conditions are present. Although the description of these conditions varies depending on which grammar book of English is consulted, the rule for final consonant doubling in the case of -ed addition may be summarized as follows: “A final consonant letter is doubled when it is preceded by one stressed vowel letter” (Koning & van der Voort, 1997, p. 33). This rule explains why verbs such as admit, prefer, rob and stop have the regular past and past-participle verb inflections admitted, preferred, robbed and stopped, which all display consonant doubling. The reason for doubling the final consonants is that these verbs are either monosyllabic (and consequently have final stress on the only vowel in the verb) (e.g., rob, stop) or that they are polysyllabic with a stressed vowel letter preceding the final consonant doubling. The reason for not doubling the final consonant is that these verbs all end in a final consonant but the consonant is preceded by either an unstressed vowel letter (e.g., answer, differ, envelop) or two consecutive vowel letters (e.g., defeat). Although the spelling rule of final consonant doubling is relatively straightforward and covers a proportionately large number of verbs in present-day English, there are some (extremely common) exceptions, which need to be taken into account when producing accurate past and past-participle verb inflections. For example, the verbs label, marvel, signal and travel all have base forms which end in a final consonant (-l) and which do not carry any final stress. Consequently, the conditions for doubling final consonants listed above are not satisfied. In keeping with the general rule listed above, past and past-participle verb inflections without consonant doubling are produced by adding -ed to the base forms of the verbs, resulting in the past and past-participle verb inflections labeled, marveled, signaled and traveled. These forms are orthographically accurate and thus grammatically possible but they are not commonly found in British English. They are the preferred forms in American English. British English, however, prefers to treat these verb inflections as exceptions and systematically chooses to double the final consonant in these cases (labelled, marvelled, signalled, travelled), which is not in keeping with the general rule. Another exception are verbs ending in -ic (e.g., frolic, mimic, picnic, traffic). For reasons related to pronunciation, they generally form their past and past-
participle verb inflections by first adding -k (and not by doubling the final -c) to the base forms and then the regular inflectional suffix -ed, resulting in the past and past-participle forms frolicked, mimicked, picnicked and trafficked. There are also final consonants letters which would seem to be found in base forms which satisfy the conditions for final consonant doubling above but which are generally never doubled (e.g., -w (allow/allowed, follow/followed, thaw/thawed), -y (play/played, stay/stayed, toy/toyed), -x (fax/faxed, fix/fixed, mix/mixed)). In addition to these more generalizable exceptions, there are also isolated cases of orthographically ‘quirky’ past and past-participle verb inflections among the regular verbs (e.g., bias/biased versus biassed, focus/focused versus focussed). It should be stressed that there is at times contradiction between various English grammar books on the issue of final consonant doubling, with some grammar books listing possible spellings which others do not regard as grammatical (e.g., benefit/benefited versus benefitted, program(me)/programmed versus programmed, worship/worshipped versus worshipped).

-e deletion. The second general spelling rule is the rule of -e deletion. Although this rule is not always consistently referred to in English grammar books, resulting in different descriptions of the exact same phenomenon (see, for example, Alexander, 1988, Biber et al., 1999, Declerck, 2003, Huddleston & Pullum, 2005), a decision was taken to describe the rule of -e deletion here in as pedagogically unambiguous a way as possible. This decision was the result of the clear focus on both relative complexity and the instructional setting in which the Dutch-speaking ESL learners participated in the experiments. As such, the rule of -e deletion may be stated as follows: If the base form of the verb already ends in -e, the regular past and past-participle verb inflections are formed by adding -d (and not -ed) to the base form of the verb (e.g., arrive/arrived, canoe/canoed, dye/dyed, free/freed, smile/smiled).46

Final -y replacement. The third and last general spelling rule is the rule of final -y replacement. Verbs that end in a consonant letter which is followed by -y generally replace the -y with an -i before adding the regular past and past-participle suffix -ed (e.g., copy/copied, envy/envied, try/tried). The presence of a consonant letter before the -y is of crucial importance here. If no consonant letter is available, the regular inflectional suffix -ed is simply added without any general spelling changes (e.g., play/played, sway/swayed, toy/toyed).

The general rules for the formation of past and past-participle verb inflections—including orthographical changes induced as a result of the three general spelling rules described above—cover the majority of verbs in present-day English but not all of them. A reference was already made above to a small group of regular verbs which display irregularities which cannot be explained using any rule available (e.g., bias/biased versus biassed, focus/focused versus focussed, worship/worshipped versus worshipped). In addition to these few regular verbs, present-
day English also has a relatively small group of irregular verbs, which display irregularities in the formation of past and/or past-participle verb inflections (e.g., begin/began/begun, drive/drove/drawn, fall/fell/fallen, go/went/gone, put/put/put). References to irregular verbs generally refrain from listing an exact number of irregular verbs. This is probably due to the dynamic nature of language and a constant flux in the exact number of irregular verbs. Instead, estimates are generally provided. The estimate of approximately 200 irregular verbs appears to be reasonably popular in publications dealing—partially or completely—with the English tense system, albeit with slight variations such as “about 200” (Biber et al., 1999, p. 394), “altogether some 200” (Dekeyser et al., 1999, p. 38) and “only about 200”. However, other scholars appear to put the estimate slightly lower or higher and are quoted using “164” (Pinker, 1999, p. 91), “about 180” (Marcus, 2002, p. 154) and “300 or so” (Teschner & Evans, 2007, p. 32). It should be mentioned that it is not altogether clear how these numbers are calculated, that is, which criteria are used to define the requirements for membership of this relatively small group of irregular verbs. Criteria may be derived based on the formal variation that many irregular verbs display. However, even then it is not always clear whether, for example, (common) irregular prefixed verbs (e.g., misread, outdo, rewrite, undergo) are systematically included in these numbers or not.

Regardless of the problematic nature of counting the number of irregular verbs, there are several statements about irregular verbs which can be made. Firstly, irregular verbs are a relatively small group of verbs (compared with the vast number of regular verbs in present-day English), with many irregular verbs occurring relatively frequently in both speech and writing. Secondly, irregular verbs are generally considered to be a closed group of verbs. Thirdly, even though irregular verbs are said to be irregular, they do sometimes show signs of predictable behavioural patterns. As a result, they are often discussed with respect to membership of a specific class or group of irregular verbs. And finally, irregular verbs may be described as irregular but it would be highly presumptive to automatically equate irregularity with (grammatical) complexity since it has been shown that other factors (e.g., frequency, (formal) salience) may influence features of (grammatical) complexity (DeKeyser, 2005).

I would like to take some time to look at the third and fourth statement about irregular verbs in more detail. The idea of irregular patterns of behaviour is one which is commonly found in English grammar books (Biber et al., 1999; Dekeyser et al., 1999; Greenbaum, 2000). As a general rule, it can be said that irregular past and past-participle verb inflections are formed in present-day English using one or more of the following four morphological processes: (1)
ablauting (e.g., drive/drove/driven, run/ran/run), (2) alveolar suffixation (e.g., bend/bent/bent, learn/learnt/learnt), (3) consonantal alternation (e.g., leave/left/left, teach/taught/taught) and (4) suppletive forms (e.g., be/was, were/been, go/went/gone) (Dekeyser et al., 1999). One of the characteristics of the past in present-day English is that it generally shows no formal paradigmatic variation in terms of number and person (i.e., conjugation), resulting in highly uniform verbal paradigms for the past (for both regular and irregular verbs). The only (common) exception to this rule is the irregular verb be, which has two forms in its verbal paradigm for the past (was/were). Although the present perfect shows a higher degree of formal paradigmatic variation than the past, the degree of variation is relative and limited to the conjugation of the auxiliary have in the present, resulting in only two morphologically different forms in the present-perfect paradigm (has/have + V-en).

It should be noted that membership of a specific class or group of irregular verbs is not exclusive and a variety of factors may influence any preference—or at times even necessity—for regular or irregular forms. It has already been stated that it is sometimes impossible to make discrete distinctions between form-related, meaning-related and use-related complexity. The semantics of the individual verb itself are a clear example of how form and meaning may interact and how the semantics of the verb itself may play a decisive role in deciding whether to use a regular or an irregular form. The verb ring, for example, is often listed as an irregular verb with the irregular past and past-participle verb inflections rang and rung respectively. In many meanings of the verb, these verb inflections are grammatically correct (e.g., He rang the doorbell upon arrival, Have you rung your friend yet?). However, when the verb ring is used in, for example, its meanings of surrounding something or someone (e.g., The police ringed the house), of putting a metal ring around a bird’s leg (e.g., All the pigeons have been ringed) or of putting a circle around something (e.g., They ringed the date on the calendar), it is used as a regular (denominal) verb. The semantics of the context should normally provide the information required to guide language users in their choices of verb inflections.

49 The past form of the verb learn may be both regular (learnt) and irregular (learnt). Several irregular verbs in present-day English have regular past and past-participle verb inflections too (e.g., dive - dove/dived - has/have dived, lean - leant/leaned - has/have leant/leaned, sneak - snuck/sneaked - has/have snuck/sneaked). Any preference for regular or irregular verb inflections may be influenced by one specific factor or a variety of factors (e.g., grammatical use, the individual verb itself, mode, regional variety, register, semantics) (Bandi-Rao, 2009; Biber et al., 1999).

50 The -en form is used in English linguistics to refer to the past-participle forms of verbs. However, this does not necessarily mean that the past participle forms of verbs always end in -en. Regular verbs in present-day English have past-participle forms which end in -ed. Thus, the past-participle forms are identical with the past forms of the same verbs. Consequently, both endings (-ed and -en) are sometimes used interchangeably to refer to this verb form (Crystal, 2003).

51 For a more detailed discussion of possible (experimental) factors see, for example, Bandi-Rao, 2009.
By way of conclusion, it can be said that even though past and past-participle verb inflections may cause formal problems for ESL learners, the majority of verbs in present-day English are fairly regular and their past and past-participle verb inflections are formed in relatively predictable ways. As far as regular spelling changes and instances of irregularity are concerned, it can be said that verb inflections may cause formal problems and, consequently, represent challenges with respect to form-related complexity. However, two issues must be stressed in this respect. Firstly, many of the formal complexities are detectable only in the written mode (e.g., final consonant doubling, final -e deletion). Secondly, there is no straightforward relationship between, on the one hand, regular spelling changes and instances of irregularity and, on the other hand, (increased) form-related complexity. Other factors (e.g., frequency, (formal) salience) may also play a role in determining overall form-related complexity. For example, the verb be is considered by most grammar books of English a highly irregular verb but its suppletive past and past-participle verb inflections (was/were and been respectively) are generally considered high-frequency verb inflections in present-day English. Consequently, many—if not most—ESL learners do not experience these verb inflections as problematic.

### 2.4.4 Meaning-related and use-related complexity

A complete linguistic account of the past and present perfect in present-day English (and of the mapping complexities that ESL learners encounter when acquiring these two tenses) must try to capture not only the formal essence of these two tenses but also their meaning-related and use-related cores, which are described as inherently more challenging in many grammars of English. In addition, such an account should shed light on possible boundaries between these two semantic neighbours in the English tense system.

Bearing in mind the adopted ternary grammar framework (form, meaning, use), the logical continuation of this chapter would dictate a discussion of the meanings and uses of these two tenses. However, the temporal meaning/use distinction in the English verb phrase is an intricate one to discern (Celce-Murcia & Larsen-Freeman, 1999) since many aspects related to meaning and use features and the boundaries of these tenses are intertwined. So much so that separate discussions of meaning and use would unnecessarily complicate matters even further. The remainder of this section will look at the actual meaning- and use-related complexities involved in choosing between the past and the present to locate bygone situations in present-day English. Since these complexities have already been extensively discussed from a largely crosslin-

---

52 The notion of frequency also plays a role, for example, in the (non-)use of uncommon irregular verb inflections by adult native speakers of English. Pinker (1994) highlights the phenomenon of regularizing lower-frequency irregular verb inflections. Some examples of common regularized past verb inflections in contemporary English are slain/slayed (instead of slew), strive/strived (instead of strove), tread/treaded (instead of trod).
guistic point of view in Chapter 1, the reader is advised to revisit this section for a detailed discussion of the complexities that Dutch-speaking ESL learners face when acquiring these L2 target features in instructional settings. In Section 2.4.3, reference was made to the fact that most discussions of verb-phrase temporality in present-day English in (pedagogical) grammars generally focus on meaning-related and use-related issues and discuss these in greater detail than form-related issues. My own concern with meaning and use is reflected in the focus on meaning and use in both the more theoretical chapters in this doctoral dissertation and in the experiments which have been carried out.53

In view of what has been said in Section 1.4 and of the remarks made in this section, it can be said that the meaning-related and use-related complexities for the past and the present perfect in present-day English are not found in all the meanings and in all the uses of both tenses. The L2 target features under investigation in this doctoral dissertation are the past and the present perfect when used to locate bygone situations in present-day English. One of the clearest meaning-related and use-related complexities for Dutch-speaking ESL learners is choosing between these two tenses to express temporal relations. For example, Dutch-speaking ESL learners are often faced with problems when having to choose between the past and the present perfect in past-zone contexts since the English and Dutch tense systems display many similarities but also some fundamental differences when locating bygone situations in such contexts. Whereas standard, formal present-day English does not allow the use of the present perfect in past-zone contexts (e.g., *I have seen her yesterday), standard, formal present-day Dutch very often does and even prefers the present perfect in those situations (e.g., Ik heb haar gisteren gezien).

2.4.5 Mapping-related complexity
When analysing tense-related learner data, it is clear that Dutch-speaking ESL learners—like many other ESL learners—are susceptible to producing a variety of ungrammatical temporal FMU mappings. The ungrammaticality may be form-related, meaning-related and/or use-related. Often, the literal Dutch translations of many of the ungrammatical temporal FMU mappings in English are perfectly grammatical in standard, formal present-day Dutch. Consequently, many lay people invoke negative L1 transfer as the obvious cause of such ungrammatical temporal FMU mappings in present-day English. However, not all of the ungrammatical temporal data can be traced to L1 transfer. Assuming that form-related complexity is not as challenging as meaning-related and use-related complexity for the temporal FMU mappings under investigation, what factors are possibly at work as far as mapping-related complexity is concerned?

---

53 For more detailed information on the form-related, meaning-related and use-related coding and scoring principles see Chapters 4, 5 and 6.
Some examples of ungrammatical FMU mappings in present-day English may be found in some the following sentences:

(2.3) (a) *He has visited his friends yesterday.
(b) Hij *beet gisteren zijn vrienden bezocht.
(c) He visited his friends yesterday.

(2.4) (a) *I know him now for ten years.
(b) Ik ken hem nu al tien jaar.
(c) I have known him now for ten years.

(2.5) (a) *Next year they are married for fifteen years.
(b) Volgend jaar zijn ze vijftien jaar getrouwd.
(c) Next year they will have been married for fifteen years.

(2.6) (a) *I will talk to him about it when I will see him.
(b) Ik zal er met hem over praten wanneer ik hem zie.
(c) I will talk to him about it when I see him.

After reading sentences (2.3a), (2.4a) and (2.5a) above, many lay people would instinctively adduce the ungrammatical English sentences to L1 transfer and the grammatical Dutch translations (2.3b), (2.4b) and (2.5b) would appear to support their claims. However, sentence (2.6a) does not support the claim that L1 transfer is the only cause of the ungrammatical English sentences. L1 transfer may play a role in the acquisition and instruction of complex, temporal FMU mappings but it is definitely not the only factor in this intricate process. The question that consequently arises is a question not about one complexity-inducing factor but rather about a possible combination of complexity-inducing factors that influence the production of such ungrammatical temporal FMU mappings by Dutch-speaking ESL learners.

In the ternary approach to grammar adopted in this dissertation for the description of the complex L2 target features under investigation, various complexity-inducing factors related to the three individual aspects of the complex mapping have already been highlighted. Form-related complexity, meaning-related complexity and use-related complexity have already been discussed. However, a more complete picture of the degree of relative, temporal SLA verb-phrase complexity present in the L2 target features under investigation must also take into account the relationship(s) between form, meaning and use. In other words, a more complete picture must also take into account the actual form–meaning–use mapping, on the one hand, and, as DeKeyser (2005) puts it, “the psycholinguistic difficulty of acquisition, that is, the difficulty of grasping the form-meaning relationship while processing a sentence in the L2” (p. 3) on the other hand. Once again, we will adopt DeKeyser’s binary approach (form, meaning) and make the necessary changes to accommodate our ternary approach (form, meaning, use).
DeKeyser (2005) refers to the transparency of the (form-meaning) relationship and to at least three factors which may cause a reduction in or lack of transparency: (1) redundancy, (2) optionality and (3) opacity. In addition to these three factors, DeKeyser (2005) also lists frequency as a factor which plays a pivotal role in determining the ease or difficulty of acquiring mappings. Let us have a look at these four factors and try to link up each factor with the L2 target features under investigation.

**Redundancy.** Temporality in present-day English can be expressed in various ways and the linguistic devices at the disposal of ESL learners are of a pragmatic nature (e.g., chronological order, scaffolding), a lexical nature (e.g., adverbials) and a morphological nature (e.g., tense). In choosing one or several of these means, ESL learners must make conscious decisions. These decisions not only reflect the actual temporal relations to be conveyed but they also take into account the semantic compatibility of the means being used. In other words, whatever means of expressing temporality are used, ESL learners must make conscious decisions to ensure that semantic alignment is achieved. Semantic alignment features prominently in Declerck's theory of tense, which states that the use of a specific tense "is wholly determined by its semantics (= temporal structure), which has to fit in with the temporal information given by the time-specifying adverbials or by the context" (p. 599). In practice, this means that (Dutch-speaking) ESL learners may often resort to using various means simultaneously in an attempt to express and reinforce temporal relations in grammatically acceptable ways. The feature of redundancy comes into play in the discussion of complexity since temporality may be conveyed by means of various (linguistic) devices which are used simultaneously in (extended) discourse. A common way of expressing, for example, past-zone temporality is the combination of tense (i.e., a morphological means) and adverbials in the linguistic environment (i.e., a lexical means).

From a psycholinguistic point of view, the co-occurrence of various devices may have implications with regard to, for example, the processing of (redundant) L2 data. The idea of processing costs was briefly referred to in Section 2.3.3, in which it was mentioned as one of several determinants of the cognitive difficulty of specific L2 language features. However, how do processing features play a role in the mapping, acquisition and instruction of temporal (verb-phrase) morphology?

A model that has been extremely influential in sketching and fine-tuning the strategies and mechanisms that (foreign) language learners use to establish

---

54 See Bardovi-Harlig (2000) and Chapter 3 in this doctoral dissertation for more details on the three means available to express temporality.
55 The term *semantic alignment* is not explicitly used by Declerck (1991, 2003, 2006) but is used in this dissertation to refer to the feature of semantic compatibility.
56 The pragmatic means of expressing temporality may also be used to express past-zone temporality but will not feature prominently in the further discussion in this chapter.
FMU mappings is VanPatten’s model of input processing (IP) (1996, 2005, 2007), which consists of two main principles and several subprinciples. Of special interest to the study of (temporal) redundancy is the first principle of VanPatten’s model, the primacy of meaning principle, and its so-called preference-for-nonredundancy subprinciple (see Figure 2.3).

---

**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.

---

In a 2007 publication, VanPatten revised the lexical preference principle and formulated the revision as follows: “If grammatical forms express a meaning that can also be encoded lexically (i.e., that grammatical marker is redundant), then learners will not initially process those grammatical forms until they have lexical forms to which they can match them” (p. 118).
Let us have a closer look at how VanPatten’s primacy of meaning principle and its corollaries fit into the broader picture of (temporal) redundancy and mapping complexity with respect to the L2 target features under investigation. When locating bygone situations in English, ESL learners may be forced to choose between either the past or the present perfect. One of the problems that they grapple with is the choice between these two tenses in past-zone contexts. How can this problem be related to VanPatten’s model of input processing in general and to the relevant principles in particular? To exemplify the problem, I will use variations of example sentences, which were discussed above. The variations are the following three sentences:

\[
\begin{align*}
(2.7) & \quad \text{(a) } *\text{She has sent me a text message three days ago.} \\
& \quad \text{(b) Ze heeft me drie dagen geleden een sms'je gestuurd.} \\
& \quad \text{(c) She sent me a text message three days ago.}
\end{align*}
\]

In (2.7c), the grammatically correct sentence in standard, formal present-day English, the past-zone context is created by a combination of lexical and morphological means. The lexical element is the adverbial *three days ago*, which can be categorized as a past-zone adverbial (see Section 1.4.1). The morphological element is *sent*, the past verb inflection of the irregular verb *send*. Semantic alignment between both elements is an absolute requirement and is present in (2.7c). Both elements are past-zone elements which are perfectly compatible in standard, formal present-day English.

As far as the processing of these elements at a sentential level is concerned, we can try to apply VanPatten’s primacy of meaning principle and its subprinciples. The primacy of meaning principle implies that ESL learners process the past-zone input for meaning before they process it for form. In keeping with VanPatten’s model of input processing, I am assuming, of course, that the ESL learners have perceived and noticed the past-zone input. Without perception and noticing processing cannot take place (VanPatten, 2004a).

Following principles 1a and 1b, the ESL learner processes the adverbial *three days ago* in the past-zone input before the past verb inflection *sent*. What is more, the ESL learner will tend to rely on the adverbial to retrieve meaning since both the adverbial and the past verb inflection encode the same past-zone meaning. The grammatical form *sent* may be deemed redundant in the past-zone input since the past-zone meaning is also encoded in the adverbial *three days ago*. This redundancy may lead to a situation in which the ESL learner will not process the grammatical form or will process it only partially (principle 1c). Caution should be exercised at this point since no processing (or partial processing) of the grammatical form does not mean that the ESL learner will not perceive and/or notice the grammatical form. The grammatical form may simply not be processed (completely) even though the ESL learner will perceive and/or notice it. The grammatical form *sent* is meaningful so according to principle 1d of VanPatten’s primacy of meaning principle the form is a likely candidate for
processing but processing may be influenced as a result of the other subprinciples. Following principle 1e, we can state that the past verb inflection *sent* is a redundant meaningful grammatical form and that the overall sentential meaning must not drain the L2 learner’s available processing resources (whatever these may be). Principle 1f is not terribly explicit in that it does not state the exact position of items in the processing of the input. The grammatical form *sent* in our example sentence comes before the lexical form *three days ago*, which is found in final position, but *sent* does not appear in absolute initial position since the subject pronoun has taken up that position. VanPatten’s sixth subprinciple is not completely transparent with respect to the consequences of these positions of forms.

If we assume that VanPatten’s primacy of meaning principle is an accurate reflection of actual SLA processes, the overall conclusion with regard to the processing of the grammatical form *sent* is that the ESL learners’ processing resources are drawn towards the lexical form *three days ago*. This does not mean that the grammatical form *sent* is not processed at all but it is not an absolute priority for ESL learners because of a variety of reasons (e.g., primacy of content words, reliance on lexical items, redundancy, available resources).

### Optionality

DeKeyser (2005) refers to the aspect of optionality and cites examples such as null subjects (in Spanish and Italian) and case marking (in Korean). He talks about the “alternating presence or absence in the presence of the same meaning” (p. 8). Even though DeKeyser appears to accept, for example, that pro-drop languages such as Spanish and Italian may or may not express subject pronouns overtly without any difference in meaning, often the use of subject pronouns does lead to a nuanced shift in meaning in these languages since overt use of such pronouns may lead to the expression of, for example, emphasis. For the sake of the current discussion, DeKeyser's idea of optionality will be referred to as optionality with semantic equivalence even though strict semantic equivalence may not be consistently present. How can optionality be related to the English tense system and to the past/present perfect distinction under investigation? Even though the English tense system has semantically demarcated verb inflections to express temporal relations, the choice between the past and the present perfect when used to locate bygone situations is often a choice between two closely related semantic neighbours in the English tense system with many instances in which both tenses are grammatically possible.

---

VanPatten’s primacy of principle meaning states that L2 learners will tend to rely on lexical items as opposed to grammatical form to retrieve meaning when both encode the exact same meaning. However, no explicit references are made to the proximity of these lexical items to the grammatical forms. At a sentential level, one can assume relatively close proximity between the lexical items and the grammatical forms. However, temporal relations in (extended) discourse are often established between lexical items and grammatical forms which are not necessarily found in close proximity. VanPatten does not express any possible correlation between changes in the proximity of these elements and changes in the way that these elements are consequently processed.
Lewis (1986) explicitly refers to this general grammatical idea of choice using the wording “grammar as choice” (p. 42). Leech (2004) also discusses this element of choice when elaborating on the choice between the past and the present perfect to refer to the expression of past time and what Declerck calls bygone situations. Declerck (2003, 2006) makes similar references to this aspect of the English tense system. In the absence of, for example, any explicit adverbials or past-zone contexts, English often allows the use of either a past or a present perfect. However, the feature of optionality in the choice between the past and the present perfect is not synonymous with DeKeyser’s concept of optionality (with semantic equivalence) since the past puts the temporal focus on THEN and the present perfect on NOW (e.g., I saw her and I have seen her). In essence, this concept of optionality is optionality without semantic equivalence. In other words, I clearly see DeKeyser’s point when he addresses optionality but optionality with semantic equivalence is not the norm in language. There is indeed often an aspect of optionality involved in the selection of grammatically correct tense forms when choosing between the past and the present perfect to locate bygone situations. However, this does not mean that the tense forms from which the selection is made are necessarily equivalents as far as meaning and use are concerned. Although not synonymous with DeKeyser’s concept of optionality referred to above, the feature of optionality or choice is a feature which has to be considered since the idea of choice may cause problems for ESL learners, who often regard grammatical choices as either choices with only one correct option or as choices with equivalent meanings and uses. Such problems are exacerbated when the choice is one between semantically closely related neighbours such as the past and the present perfect in present-day English. The SLA challenge then consists of fine-tuning semantic nuances, which is an intricate and at times conceptually abstract process for many L2 learners.

**Opacity.** DeKeyser (2005) refers to opacity as “a complex form of the problem of low form-meaning correlation” (p. 8). The obligatory use of the past in past-zone contexts and the ungrammatical use of the present perfect in past-zone contexts may be interpreted as a choice between two mappings which are not terribly opaque in past-zone contexts for many ESL learners. However, focusing on only the two tenses involved in this choice is only one part of a much larger challenge. Reference has already been made to the fact that tenses are used based on their semantics and that semantic alignment plays a vital role in helping ESL learners process tenses accurately, meaningfully and appropriately. This is also the case when choosing between the past and the present perfect in, for example, past-zone contexts. An ideal—albeit unrealistic—situation would be a scenario in which tenses represented unique temporal FMU mappings, that is, one tense form with one tense meaning and one tense use. Sadly, this is not the case in present-day English (nor in most of the other known natural languages). Tenses are often used with various meanings,
in various ways and with only nuanced semantic differences at times. Often, the temporal FMU mappings are the result of the inherent semantic properties of the tenses combined with contextual features. In addition, the past and the present perfect are two tenses of a much more elaborate tense system in present-day English. Not only does present-day English have more than two tenses, with every tense reflecting inherent semantic properties (Declerck, 1991, 2003, 2006), the English language—like many other natural languages—also has a tendency to combine tense with other grammatical categories in the verb phrase (e.g., aspect, mood). In other words, the grammatical category of tense casts its net relatively wide and even though the choice between the past and the present perfect in past-zone contexts may not be opaque from an absolute point of view, the net within which this choice is to be made may contribute to an even higher degree of opacity and thus to a higher lack of (temporal) transparency for ESL learners.

**Frequency.** The role of frequency in language learning is an intricate one, which was already highlighted cursorily in Section 2.4.3, where it was mentioned with respect to form-related complexity. However, the role of frequency may be said to extend beyond mere formal features. It may also be applied to the other aspects of the temporal FMU mappings under investigation: (1) meaning and (2) use.

Initial SLA investigations into the role of frequency were focused on the relationship between input frequency and the order of acquisition (Ellis, 2008). The frequency-focused theory (or hypothesis) which guided early SLA studies and which was formulated in the 1970s became known in SLA research as the frequency hypothesis. In its simplest form, the hypothesis stated that the order of L2 acquisition was determined by the frequency of linguistic items in the input and that high-frequency linguistic items would be acquired before low-frequency linguistic items. Many SLA scholars decided at the time to test the claim(s) expressed in the frequency hypothesis but the results from the subsequent studies were not conclusive. Some SLA studies showed significant effects for the research variable input frequency (e.g., Larsen-Freeman, 1976a, 1976b; Lightbown, 1983) whereas others did not find any pronounced correlations between input frequency and accuracy (e.g., Long & Sato, 1984; Snow & Hoefnagel-Höhle, 1982). Later SLA studies and reviews have been able to highlight the role of frequency on SLA learning more comprehensively and more convincingly (e.g., Gass & Lakshmanan, 1991; Goldschneider & DeKeyser, 2001). The issue that the majority of the later SLA studies into the effects of input frequency have reported though is related to the unique effects of input frequency. Frequency has generally been recognized as an input-related factor which plays a role in the multifaceted L2 acquisition process. However, SLA researchers have also been explicit about stressing the actual role of input frequency. The overall consensus in contemporary SLA research is that input frequency is one of several possible determinants and that it is part of a more
complex set of interactions in the L2 acquisition process. In their 2001 study, Goldschneider and DeKeyser highlight the results of their meta-analysis with respect to the effects of five determinants which they selected for discussion: (1) perceptual salience, (2) semantic complexity, (3) morphophonological regularity, (4) syntactic category and (5) frequency (in the input). They acknowledge that other determinants possibly exist and that not all of the determinants are necessarily input-related determinants. They explicitly list L1 transfer as a determinant which is not input-related but there are others (e.g., processing constraints as stated, for example, by VanPatten, 1996, 2005, 2007). However, since the focus of their study was on the effects of the properties of the so-called grammatical functors, they made a conscious decision not to investigate the effects of possible determinants external to the functors. As part of their conclusions and implications section Goldschneider and DeKeyser (2001) state the following:

It would also be interesting to try to tease apart the individual determinants’ effects on acquisition in order to establish with more certainty whether the combination of factors accounts for the order through a cumulative effect or through the interaction of the factors. (p. 38)

For the temporal FMU mappings under investigation in this doctoral dissertation, input frequency no doubt plays a role in the L2 acquisition process. As such, frequency may affect all of the three aspects of mappings (form, meaning, use) and, of course, the actual acquisition (and instruction) of the temporal FMU mappings by ESL learners (and teachers) in instructionally explicit settings. The mappings themselves (i.e., the past and the present perfect in present-day English to locate bygone situations) may not be completely straightforward for Dutch-speaking ESL learners. However, they cannot be considered altogether obscure since they do in fact appear relatively frequently in the input to which the participants in the studies are exposed inside and outside the respective instructional settings. The reader should simply keep in mind either possible cumulative effects or the interaction of determinants (referred to above by Goldschneider and DeKeyser (2001)) in the L2 acquisition and instruction of the complex, temporal FMU mappings under investigation.

2.4.6 Crosslinguistic influence and outcome measure complexity as complexity-inducing factors

The aim of this section is to shed light on two other complexity-inducing factors which have already been referred to fragmentarily in this chapter but which merit further discussion within a context of SLA complexity in instructionally

---

Goldschneider and DeKeyser (2001) refer to the grammatical input using the term grammatical functors, which they use as a less theory-burdened synonym for grammatical units or grammatical morphemes.
explicit settings: (1) crosslinguistic influence (CLI) and (2) outcome measure complexity.

**Crosslinguistic influence (CLI).** The first complexity-inducing factor to be discussed in this section is related to an issue which was already raised in the previous sections on complexity and which is generally referred to in contemporary SLA discourse as crosslinguistic influence (CLI) or transfer. The concept of CLI has been the object of investigation for quite some time in a variety of linguistic subdomains (e.g., contact linguistics, language typology, language universals) (Odlin, 2003). The linguistic subdomain of SLA is no exception and has investigated CLI since its inception. The aim of this section is not to provide a detailed account of CLI by means of, for example, exhaustive lists of SLA studies carried out on the topic. Rather, the aim in this section is to highlight the concept of CLI and how it may contribute to relative, temporal SLA verb–phrase complexity in experimental studies.

CLI is a feature which is generally considered to be present when several languages come into contact. The phenomenon of ‘coming into contact’ may also be referred to as interlingual or crosslinguistic contact. The expression ‘come into contact’, however, can take on a variety of meanings depending on the exact context in which it is used. For the SLA complexity-related discussion at hand, I will assume that interlingual contact is the coming into contact of two languages, which takes place—psycholinguistically—in the L2 learners’ heads when they are asked to call upon their multi-competence of two (or possibly more than two) languages. The two languages under investigation here are the L2 learners’ native language (i.e., Dutch) and their target language (i.e., English), which are typologically closely related languages.

The initial contrastive SLA research agenda in the mid-twentieth century relied heavily on contrastive comparisons between two languages and on predictions which were dependent on the comparisons undertaken (Brown, 2000). The hypothesis which resulted from such comparisons in the 1950s and 1960s was generally referred to as the contrastive analysis hypothesis (CAH) (Lado, 1957). As far as acquiring a second language was concerned, the CAH claimed that the intractable problem for L2 learners in the SLA process was the negative transfer which took place from the language learners’ L1s to the language learners’ L2s (Brown, 2000; James, 1998; Sharwood Smith, 1994). As such, the focus of attention for CAH advocates was on the predictive value of the hy-

---

61 The terms crosslinguistic influence and transfer are the most commonly used terms to refer to the concepts of crosslinguistic influence in the field of second language research. However, other common ways of referring to crosslinguistic influence are, for example, language transfer, linguistic interference, the role of the mother tongue, native language influence and language mixing (Odlin, 2003).

62 For overviews of CLI, the reader is advised to consult more comprehensive SLA-specific accounts on the topic (e.g., Brown, 2000; Ellis, 1994, 2008; Gass & Selinker, 2001; Odlin, 1989, 2003; Ortega, 2009; Ringbom & Jarvis, 2009).

hypothesis as far as L1-induced SLA complexity/difficulty was concerned. The claims that were made using the CAH were supported by many scholars in the form of empirical methods of prediction.\footnote{See, for example, Brown, 2000, and Odlin, 2003, for discussions of the contrastive analysis hypothesis.} One such model was Stockwell, Bowen and Martin’s hierarchy of difficulty (1965), which was drawn up to make predictions about the relative difficulty of specific L2 target language features for the language pair L1 English and L2 Spanish. In discussions of CLI, Stockwell, Bowen and Martin’s oft-cited hierarchy of difficulty model is an extremely popular model which has been hotly debated and critiqued. It clearly shows that the concept of CLI (and its possible effect on the SLA process) is not a new and unexplored phenomenon (Brown, 2000; Cook, 1993; DeKeyser, 2005; Ellis, 2008; Sharwood Smith, 1994). The 1960s and 1970s, however, saw a shift in the focus of investigation from predicting L2 difficulties to the actual analysis of learner language, which was also referred to using the term interlanguage. This shift occurred after it was shown that the CAH—in both its strong and weak version—was not as tenable as had been claimed and that observations of learner language data did not validate the predictions of difficulties resulting from interlingual comparisons.\footnote{The strong version of the CAH, which was considered by many untenable, focused on the predictive value of the CAH, which was the result of contrastive analysis. The weak form of the CAH focused on observational use of the CAH rather than predictive use (Brown, 2000).} The discussion of CLI in this section should be viewed from an interlanguage point of view. By no means do I wish to provide any predictive value for the comparison of differences and similarities between temporal FMU mappings in English and in Dutch.

Throughout this chapter the focus has been on the problems that Dutch-speaking ESL learners experience when having to choose between the past and the present perfect when locating bygone situations in present-day English. These problems often result in wrong choices being made, namely, in choosing, for example, the ungrammatical present perfect to refer to bygone situations in past-zone contexts (e.g., *I have seen her yesterday* instead of the grammatically correct option *I saw her yesterday*). In other words, the past is underused and the present perfect is overused. Consequently, this would point to possible problems that Dutch-speaking ESL learners experience when mapping form to meaning and when using that mapping correctly. In fact, for many lay people negative L1 transfer is the only (intuitive) explanation that sufficiently explains these problematic mapping features. To them, Dutch-speaking ESL learners are in danger of having their language ability clouded as a result of negative L1 transfer. Possible negative L1 transfer, however, is only part of the much more complicated and larger temporal puzzle under investigation, which has been highlighted in detail in this chapter. The present perfect in present-day English is not per se a problem across all of its uses. In fact, many uses of the English present perfect such as indefinite readings (e.g., *I have never seen that man before*,
Have you spoken to her? and resultative readings (e.g., Someone has left the door open, They have already given their contribution) appear to be relatively less problematic for Dutch-speaking ESL learners. It is only when the options for locating bygone situations by means of the past or the present perfect are discussed that many contrastive grammars explicitly highlight possible problems based on the language learners’ L1s (e.g., Aarts & Wekker, 1993, De Moor, 1998; Koning & van der Voort, 1997; Mackenzie, 1997 (for Dutch), Lambotte, 1998 (for French), Hoffmann & Hoffmann, 2001, 2005; Ungerer, 2000; Ungerer et al., 2009 (for German)).

Thus, from a (purely) CLIC-inspired point of view the discrepancy between grammatical and ungrammatical uses of the English present perfect could be explained as instances of positive and negative transfer respectively. However, the discussion of complexity in this chapter has already highlighted other intractable problems with respect to the choice between the past and the present perfect to locate bygone situations in present-day English. These problems are not the unique result of pure, L1-induced transfer. They are largely the result of a combination of more universal developmental patterns of specific tense morphology acquisition, L1 transfer and possibly other complexity-inducing factors. Much more than straightforward instances of CLIC is at play in the acquisition of these complex temporal FMU mappings, which are also problematic for ESL learners whose L2s share mapping similarities with English. Thus, with respect to the acquisition of the complex, temporal FMU mappings under investigation, Dutch-speaking ESL learners experience both universal, developmental problems of acquisition and specific (L1-induced) problems of acquisition. In other words, the determinants of SLA complexity are varied and consist of interacting factors of both a universal nature and a specific nature. Even this relatively simplistic division of complexity into universal developmental complexity and L1-induced complexity is, of course, a simplified approach to the acquisition of complex, temporal FMU mappings since a whole range of other factors may interact in their own unique ways to determine the various aspects of SLA complexity (e.g., interlingual identification/psychotypology, transfer variability, task complexity). Ortega’s (2009) selection of words puts the intricate relationship between transfer and other SLA-related factors into context:

In addition, knowledge of the L1 impacts on L2 acquisition subtly and selectively, sometimes resulting in strikingly different negative and positive consequences for different learner L1 backgrounds, at different stages of development or proficiency and for different areas of the L2. (p. 31)

66 ESL learners—including Dutch-speaking ESL learners—face other temporal challenges when acquiring features of the English tense system. See, for example, Section 2.4.5 for other instances of temporal problems in the English verb phrase.
The interest in CLI in SLA discourse has remained strong (Odlin, 2003). This unabated interest is clearly visible in the pages devoted to CLI in most, if not all, of the introductory publications dealing with SLA (e.g., Brown, 2000; Ellis, 1994, 2008; Ortega, 2009; Saville-Troike, 2006) and in the more specialized publications dealing with transfer in the field of SLA (e.g., Jarvis & Odlin, 2000; Odlin, 1989, 2003; Ringbom & Jarvis, 2009). The issue of transfer will be revisited in Chapter 3, where its role will be discussed specifically in relation to the study of tense morphology.

**Outcome measure complexity.** The research that was carried out for the three studies in this doctoral dissertation is experimental in that participants were invited to take part in sessions in which they received explicit instruction dealing with the meanings and the uses of the past and the present perfect to locate bygone situations in present-day English. Data collection took place in a pretest/posttest format for which the participants received contextualized input and were subsequently asked to interact with the input in context. The four outcome measures that were used for the data collection procedure were the following: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).

As far as the degree of outcome measure complexity is concerned, it is extremely difficult—if not impossible—to draw up some definitive form of ‘objective’ outcome measure complexity hierarchy because of the multifaceted combination of the input (e.g., text length, text topics), the practice-based instruction (input practice, output practice), the outcome measures (GJ, SR, CCR, TR), other experimental conditions and subjective factors (e.g., outcome measure familiarity, time pressure, vocabulary load) found in the experimental studies in this doctoral dissertation. One concrete approach to outcome measure complexity is the process of item analysis. Calculating item difficulty indices could provide us with an overview of item difficulty data. Bachman (2004) defines item difficulty as follows:

> Item difficulty is the proportion of test takers who answered the item correctly, for R-W [right or wrong] scoring, or, for P-C [partial credit] scoring, the average score on the item. We can also calculate the difficulty or proportion of test takers who chose the different distracters, for R-W scoring, or for the different item scores, for P-C scoring.

Item difficulty is an item characteristic that is relevant to both NR [norm-referenced] and CR [criterion-referenced] tests, and will be calculated in the same way for both types of tests. (p. 122)

---

67 Bachman (1990, 2004) and Purpura (2004) provide descriptions of grammar outcome measures but refer to the measures using the term (grammar) task types. See Chapters 4, 5 and 6 for detailed information on both the nature of the outcome measures and other methodological and design-related issues in the studies in this doctoral dissertation.

68 See, for example, Bachman, 2004, and Bachman and Kunnan, 2005, for more detailed information on calculating item statistics for testing purposes.
What is clear from the definition above is that Bachman’s definition of item difficulty is strongly reliant on subjective factors since the test takers’ answers play an essential role in calculating item difficulty and, consequently, outcome measure complexity and overall test complexity. How, though, can objective factors be separated from possible subjective factors when calculating item difficulty using Bachman’s equation? No answer is provided in the literature on this intricate issue. The concept of outcome measure complexity will be revisited in Chapters 4, 5 and 6, where Studies 1, 2 and 3 will be discussed in their general setup and in their experiment-specific details.

2.5 Conclusion
The aim of this second chapter was to highlight and discuss the concept of temporal complexity in SLA contexts in an attempt to provide the reader with a more comprehensive and balanced picture of temporal complexity as defined in this doctoral dissertation. In addition, the chapter sought to motivate the choice of the L2 target feature under investigation in this doctoral dissertation. In the first section (Section 2.2), I provided an operational definition of what I have termed temporal SLA verb-phrase complexity, incorporating both the tripartite approach to grammar (form, meaning, use), which was highlighted in the introduction, and (mainly) qualitative features of various forms of complexity which have been discussed in this chapter.

Since the concept of complexity is such a multifaceted concept, it was imperative to discuss terminological and conceptual issues which were taken into account to operationalize complexity. Those issues were highlighted in the second section of this chapter (Section 2.3). The first two definitional considerations were related to the concept of linguistic complexity, which has been discussed in terms of two oppositional distinctions: (1) absolute versus relative linguistic complexity and (2) global versus local linguistic complexity. Subsequently, I turned the focus to complexity in SLA research, where it has been and still is an extremely popular concept. Generally speaking, complexity in SLA studies is operationalized as either a dependent or an independent variable. In addition to discussing this operationalization aspect, I also explained how the absolute–relative linguistic complexity distinction may be applied in SLA research. Subsequently, I highlighted in more detail the concept of complexity with respect to the acquisition of L2 mappings.

The third and final section (Section 2.4) tied up the operational definition of temporal SLA verb-phrase complexity with the discussions of the definitional features and applied the insights into the concept of complexity to the acquisition and instruction of the L2 target features under investigation in this dissertation. From this analysis it became clear that ESL learners face a variety of challenges, which are not simply form-related, meaning-related and use-related but which exceed these boundaries and operate at other levels too (e.g., the L2
mapping(s) itself (themselves), the psycholinguistic challenges that the ESL learners are faced with when trying to grasp L2 mappings).

The overall conclusion with respect to Dutch-speaking ESL learners and the L2 target features under investigation is the following: The concept of temporal SLA verb-phrase complexity is intricately interwoven with an array of factors found in both the L2 target features selected for instruction and in the SLA process itself. Not only are Dutch-speaking ESL learners faced with form-related, meaning-related and use-related instances of linguistic complexity, they are also faced with the challenge of mapping all three aspects onto L2 target features during online processing, a process which is inherently complex for many—if not most—ESL learners. The target features under investigation, the past and present perfect when used to locate bygone situations in present-day English, pose universal (linguistic and SLA-specific) challenges which are experienced by most—if not all—Dutch-speaking ESL learners at different levels of proficiency. However, the universal challenges do not provide a complete and accurate reflection of possible problems with which Dutch-speaking ESL learners grapple. In addition to the universal challenges, Dutch-speaking ESL learners also face specific challenges such as crosslinguistic influence. Add to that the specificity of the instructional settings and the picture of obvious L1 transfer as the single cause of complexity—which is generally conjured up by lay people—is no longer as accurate as it is represented. Although there are many factors which contribute to relative, temporal SLA verb-phrase complexity, the distinction between universal challenges and specific challenges is an important one since it provides a more complete picture of the complementary challenges that Dutch-speaking ESL learners face when acquiring tense.

The following chapter (Chapter 3) will investigate how the acquisition and instruction of tense (and the broader concept of temporality) have been approached in the field of SLA. Chapter 3 will show how the study of L2 temporality developed from incidental investigations into temporal morphology to methodological investigations with targeted areas of interest. In addition to describing this development, important aspects in the study of L2 temporal (e.g., developed research methodology) will be highlighted and discussed. In addition, the nature and the role of instruction will be discussed in the SLA process. Using two well-known approaches to SLA, input processing and skill acquisition theory, a comparison will be drawn with respect to the views that both theories have on the role of practice in the SLA process. Since the concept of practice was an important element of the instruction provided in the experimental research, comparing both approaches in light of their views on practice will provide us with information which is valuable in drawing up research hypotheses.
CHAPTER 3
TENSE AND INSTRUCTED SLA

Like students of art who walk around a great statue, seeing parts and aspects of it from each position, but never the whole, we must walk mentally around time, using a variety of approaches, a pandemonium of metaphor.

(Robert Grudin)

3.1 Introduction
After investigating the challenges involved in defining and operationalizing both the grammatical category of tense (Chapter 1) and complexity features related to the grammatical category of tense in linguistic and SLA contexts (Chapter 2), I would now like to place the focus of investigation on a more SLA-specific context, that is, the context of L2 temporality and instructed SLA. The focus in this chapter will be on two specific aspects of acquiring L2 temporality: (1) the ways in which L2 temporality has been and is approached in the field of SLA and (2) the ways in which instruction, more precisely, input (practice)-based and output (practice)-based form-focused instruction, may be conceptualized when investigating the effects of explicit instruction.

In this third chapter, I will investigate three issues in greater detail. The first section of this chapter (Section 3.2) will look at how the grammatical category of tense functions as one of three means (pragmatic, lexical, morphological) used to express temporality. In SLA research, one of two approaches is generally used to investigate temporality. Depending on how one approaches temporality and its key features, a choice is usually made to adopt either a form-oriented approach or a meaning-oriented approach to L2 temporality. Both approaches will be discussed in this chapter. In addition, the ways in which tense is conceptualized in both approaches will be highlighted with a clear focus on the meaning-oriented approach. The reason for focusing on the meaning-oriented approach is the experimental research carried out for this doctoral dissertation, which is reported on in Chapters 4, 5 and 6. This research is predicated on a meaning-oriented approach to L2 temporality.

In the second section of this chapter (Section 3.3), the focus will be placed on one specific form of SLA, that is, instructed SLA. The title of this dissertation makes a reference to form-focused instruction (FFI) and in the second section the reader will be guided through the conceptualization of FFI in instructed SLA. At first glance, form-focused instruction may be at odds with a meaning-oriented approach to L2 temporality but clarification of this seemingly contradictory terminological distinction will shed light on the overall approach
adopted for the experimental research carried out for this dissertation. In addition, two approaches will be presented which both offer options for practice-based instructional types: (1) input processing and (2) skill acquisition theory. In both approaches, practice plays an essential role. However, because of inherent differences between both approaches, the roles and implications of practice are not identical. Consequently, predictions with regard to L2 type-of-instruction theory and the roles of practice vary and will be explained in approach-specific contexts.

3.2 Investigating the L2 acquisition of temporal expression

3.2.1 Early investigations into temporal expression

In the first chapter of her 2000 book *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use*, Kathleen Bardovi-Harlig (2000) writes that “temporal expression, or what C. Smith (1980) has called “time talk”, has come into its own as an area of research in adult second language acquisition” (p. 1). As true as these words were when the book was published in 2000, a decade later, we have witnessed and are still witnessing an enormous increase in the number of SLA publications dealing with temporality and the acquisition of highly specific and constantly developing L2 temporal features (e.g., Ayoun & Salaberry, 2005; Bardovi-Harlig, 1999, 2001, 2006; Bardovi-Harlig & Comajoan, 2008; Collins, 2002; Salaberry, 2008; Salaberry & Shirai, 2002). L2 temporality is now anything but a neglected area of SLA.

For all the methodological systematicity that many of the past and contemporary studies into L2 temporality display, the first studies which investigated tense-aspect morphology did so but in a purely incidental manner (Bardovi-Harlig, 2000). In essence, they investigated tense-aspect morphology as part of much broader research interests. The two types of studies to do this were the morpheme (order) studies and the phonetic constraints studies. The morpheme studies, which were initiated in the 1970s (e.g., Bailey, Madden, & Krashen, 1974; Dulay & Burt, 1973, 1974; Krashen, Butler, Birnbaum, & Robertson, 1978) and stopped in the early 1980s,69 investigated a set of morphemes (e.g., articles, auxiliary *be*, irregular past, past tense -ed, plural -s, progressive -ing, third person -s) and initially sought to shed light on the order in which these morphemes were acquired by L2 learners. At a later stage, the research interest in the morpheme studies also included sequential features of the L2 acquisition of the morphemes being investigated (Ellis, 2008).70 The phonetic constraints...
studies, which were initiated in the 1980s and were still being carried out in the 1990s (e.g., Bayley, 1994; Wolfram, 1984, 1985, 1989; Wolfram & Hatfield, 1986), tried to account for the frequency with which verbs occurred with past-tense verbal morphology. The overall outcome of the phonetic constraints studies (with respect to L2 English) was that the phonetic realization of the English past tense depended on phonological environments (Bardovi-Harlig, 2000). Both the morpheme studies and the phonetic constraints studies placed the research focus on the morpheme, that is, on form. By contrast, meaning was not considered an important feature worthy of too much attention and any separate lines of investigation at the time of the morpheme studies and the phonetic constraints studies.

A shift in the research focus came about in the 1980s as investigations of the expression of temporality found their own niche in SLA research and attempts were made to disentangle (temporal) form and meaning. Whereas the 1970s focused mainly on the acquisition of temporal morphology as form, the 1980s shifted the research focus to the study of how temporal morphology was acquired as the surface realizations of underlying temporal semantics. This shift was, in essence, the result of two factors: an increased interest in the semantics of both interlanguage in general and of temporal semantics in particular (Bardovi-Harlig, 2000). The remainder of this first section will focus on two main approaches to investigating L2 temporality that have crystallized in SLA research: (1) the form-oriented approach and (2) the meaning-oriented approach. Both approaches will be discussed and compared below. The focus will be placed on the meaning-oriented approach in an attempt to explain in detail the approach adopted for the experimental research carried out for this dissertation.

3.2.2 The form-oriented approach

The form-oriented approach to investigating the expression of L2 temporality explores the distribution of (emerging) verbal morphology. It does so by regarding features of temporal verbal morphology as indicators of an underlying system of interlanguage temporal semantics. Form-oriented studies into L2 temporality shed light on how and where a specific temporal verb form is used in L2 learners’ interlanguages. Figure 3.1 visualizes the form-oriented approach morpheme studies referred to above, investigations into the order of acquisition looked at the order in which L2 learners acquire a variety of grammatical morpheme irrespective of the L2 learners’ L1s (e.g., Do ESL learners learn past tense -ed before, for example, progressive -ing?). Investigations into the sequence of acquisition looked at the sequence in which a specific grammatical morpheme is acquired (e.g., Do ESL learners learn progressive -ing in a single step or do they pass through stages on their way to mastering this target feature?).

71 Both the morpheme studies and the phonetic constraints studies had other flaws too (e.g., a focus on the end point of acquisition and not the emerging system of tense and aspect). However, for the studies in this dissertation these flaws will not be discussed in detail. For a more detailed discussion of some of the flaws, the reader is advised to consult Bardovi-Harlig, 2000.
using past tense verbal morphology, more specifically, the simple past in present-day English as a verbal form with various (temporal) meanings.

---

**Figure 3.1.** *A form-oriented approach to the simple past in present-day English*

What is clear from Figure 3.1 is that the simple past in present-day English is a verbal form which has several meanings. Consequently, it is a form which may be used in several linguistic environments. The primary use of the simple past in present-day English is its use to refer to past time as is the case in the following examples:

(3.1) *We visited her two months ago.*

(3.2) *I was surprised to see her standing there as the car drove up the road.*

(3.3) *They visited their grandmother every summer when they were young.*

As is clear from (3.1), (3.2) and (3.3), the simple past in present-day English may be used to refer to single events in the past (3.1, 3.2), in which case it is often referred to as the event past, or to a repeated event in the past (3.3), in which case it is often referred to as the habitual past (Greenbaum & Quirk, 1990). The event may take place over an extended period of time, but this need not necessarily be the case.

In addition to its primary use, the simple past in present-day English also has secondary uses, which a form-oriented approach to investigating L2 temporality would also take into account. The simple past may be used, for example, as a hypothetical past. The hypothetical past does not actually refer to past time
but to present or future time. Examples of this secondary use may be found in the following sentences:

(3.4) *I wish he knew what she has been doing for him.*

(3.5) *If I had enough money, I would help you out.*

(3.6) *Would you be happy if I sold my motorbike?*

There is another secondary use of the simple past, often referred to as the attitudinal past (Greenbaum, 2000; Greenbaum & Quirk, 1990), which is described as a more polite or more tentative alternative to refer to present states of mind. Examples of such an attitudinal use may be found in the following sentences:

(3.7) *Did you want to talk to me?*

(3.8) *I wondered whether I could quickly have a word with you about this matter?*

(3.9) *I wanted to know whether you are coming to the party.*

Although both the hypothetical past and the attitudinal past show an element of pastness (or remoteness), it is not the temporal pastness associated with the simple past in its primary use. The pastness in (3.4) to (3.9) is of a figurative nature in that the speaker creates a metaphorical distance and not a temporal distance.

A form-oriented approach to investigating L2 temporality generally takes into account various meanings and uses of a temporal verb form and looks at how these meanings and uses emerge, are distributed and develop over time in L2 learners’ interlanguages.

### 3.2.3 The meaning-oriented approach

The meaning-oriented approach to investigating L2 temporality takes as its point of departure a functional approach to L2 temporality and investigates the expression of semantic concepts through various means. In SLA research, meaning-oriented studies investigate the varied range of means that L2 learners use to express (temporal) semantics. Meaning-oriented studies provide a broad picture of investigation but not necessarily a broader picture than the form-oriented studies. As is clear from Figure 3.1, form-oriented studies cover various meanings and uses but do this with a formal common denominator: the verbal form under investigation. Meaning-oriented studies have as a common denominator a semantic concept and, subsequently, investigate the various means which may encode that semantic concept. Figure 3.2 provides a schematic representation of the approach adopted by meaning-oriented studies of
L2 temporality. It uses the semantic concept of ‘bygone-ness’ under investigation in this dissertation.

Figure 3.2. A meaning-oriented approach to the expression of ‘bygone-ness’ in present-day English

The developmental picture provided by meaning-oriented studies is reflected in three stages of L2 temporal semantics. The stages which are generally distinguished in meaning-oriented studies into L2 temporality are (1) the pragmatic stage, (2) the lexical stage and (3) the morphological stage (Bardovi-Harlig, 2000). The expression of L2 temporality exhibits a more or less fixed route from pragmatic to lexical to grammatical means (Dietrich, Klein, & Noyau, 1995; Giacalone Ramat & Banfi, 1990; Meisel, 1987), with some scholars suggesting that this three-stage acquisitional route is universal and not dependent on the languages under investigation (Giacalone Ramat & Banfi, 1990). However, agreement on the route of acquisition of L2 temporality is lacking, which is reflected in the many hypotheses about the acquisition and the development of L2 temporality (Salaberry, 2008).

Before discussing the three stages that have been referred to above, it is important to underscore what is meant with the term *stage*. Bardovi-Harlig (2000) is extremely explicit when she says that by stage of acquisition she means “a developmental period that can be characterized by the use of a particular feature” (p. 13). Bardovi-Harlig (2000) continues her discussion by stating the following about the concept of stage:

The concept of stage is used widely in the study of the acquisition of temporal semantics by a range of researchers from different theoretical backgrounds (e.g., Andersen 1986a, 1986b, 1989, 1991; Dietrich et al., 1995). For example, the meaning-oriented approach may characterize a stage by the use of adverbials and connectives to make temporal reference, and by the absence of verbal morphology. That means that the dominant means of temporal reference is lexical, but it does not mean that features from an earlier stage or a later stage are entirely absent. Likewise, a form-oriented study may describe a stage as being characterized by the use of simple past with a certain category of verbs. That means that the dominant tense-aspect morpheme is the simple past, but it does not mean that other tense-aspect morphology is never used, or that
the past is not used elsewhere. The boundaries of stages are not abrupt because the ac-
quisition of temporal expression is gradual. Nevertheless, it is possible to identify
stages in the development of interlanguage temporality by identifying the characteristic
means of expression of each stage. (p. 13)

Although the characteristic use of pragmatic, lexical, and morphological devices can be
associated with distinct stages of interlanguage development, characteristic use is not
the same as exclusive use. (p. 48)

What is clear from Bardovi-Harlig’s (2000) references to the concept of stage is
that the three stages generally recognized in the development of interlanguage
 temporality follow a chronological route of development with the pragmatic
stage as the first stage, the lexical stage as the second stage and the morphologi-
cal stage as the third stage. However, the development of L2 temporality is
both dynamic and recursive at the same time. In other words, although an L2
learner’s interlanguage temporality may have reached, for example, the lexical
stage of development of L2 temporality, it may also display features of the
pragmatic and/or the morphological stage(s). What is of importance is the de-
gree of characteristic use.

The pragmatic stage is regarded as the earliest stage of development of L2
temporal semantics. It is characterized by the absence of any systematic use of
tense-aspect morphology. Meisel (1987) and Schumann (1987) report four dis-
tinct pragmatic means of establishing temporal reference:

1. scaffolded discourse, that is, the reliance on the contributions from fel-
low speakers,
2. implicit reference, that is, the reliance on references inferred from spe-
cific contexts,
3. the reliance on contrasting events,
4. the use of chronological order.

Although the four pragmatic means above are often distinguished, L2 learners
may use various pragmatic means simultaneously.

The lexical stage is regarded as the second stage and is characterized by a va-
riety of lexical means. The most common examples of such lexical means are
the following:

1. temporal adverbials (e.g., \textit{at 9 o’clock in the morning, yesterday, then, several
months ago, last week, 20th of June}),\footnote{It is generally recognized that temporal adverbials consist of a heterogeneous group of adver-
biales and may be subdivided into semantic types such as adverbials of position, adverbials of
duration, adverbials of frequency and adverbials of contrast, which are acquired in stages. For a
more detailed discussion of the division of temporal adverbials and their sequenced acquisition,
see, for example, Klein, 1993, 1994, and, Noor, 1993.}
2. locative adverbials (e.g., here, in Spain, at my sister’s, there),
3. connectives (e.g., first, (and) then),
4. verbs (e.g., begin, commence, end, finish, start, stop)

Many descriptive and pedagogical grammars of English highlight the interaction between the various means used to express temporality in English, especially the interaction between lexical means and morphological means. For all the attention that has been and is still being dedicated to lexical means of expressing temporality in grammars of present-day English, it is the research carried out by scholars investigating adult SLA that has shed light on the primacy of non-morphological means of expressing temporality (Bardovi-Harlig, 2000). At the lexical stage, verbs often occur in morphologically invariant forms (e.g., eat, greet, swim in present-day English), which may be either morphologically unmarked forms (e.g., base forms in present-day English) or morphologically relatively simple forms (e.g., third person singular present forms in Romance languages), which are generalized to non-targetlike contexts (Andersen, 1991; Bardovi-Harlig, 1995; Bardovi-Harlig & Reynolds, 1995; Giacalone Ramat & Banfi, 1990). Applied linguists generally agree that lexical means of temporal expression are important and that this importance is reflected in the functional load of lexical means of temporal expression in learner production. However, in Chapter 2 a reference was already made to possible processing problems that (Dutch-speaking) ESL learners may experience when lexical means and morphological means of temporal expression interact. Studies into input processing suggest that lexical means are so prominent that they often take precedence over morphological means when L2 learners process input. This finding would suggest that the heavy negative effects resulting from the large functional load of lexical means may have to be ‘counterbalanced’ when pedagogical considerations are made for tense-related instructional intervention.

The third stage of the development of L2 temporal semantics—the morphological stage—is the last stage that L2 learners enter into. Typical of the morphological stage is that it consists of verbal morphology which is initially extremely unstable and unsystematic. In addition, temporal verb morphology requires a relatively long time to stabilize. There are various reasons for the slow and gradual development of temporal verb morphology. Firstly, it is generally accepted that temporality is an intricate network of multifarious relationships, which are determined by a variety of variables. Consequently, acquiring verbal morphology is challenging in that L2 learners have to acquire a series of temporal rules—either implicitly or explicitly—which govern the forms, meanings and uses of temporal verb forms and which are not necessarily always

---

73 The intricate interaction between the grammatical category of tense and, for example, temporal adverbials has already been highlighted in Chapter 1.
74 In present-day English, morphologically unmarked forms are also referred to as base forms or default forms.
restricted to purely temporal features (e.g., interaction between tense and aspect, tense and mood, pragmatic, lexical and morphological means). The sheer intricacy inherent in temporality results in slow and gradual development of lower-level stages before higher-level accuracy, meaningfulness and appropriateness are achieved. Secondly, L2 learners appear to rely heavily on lexical means of temporality when processing L2 input. In turn, this reliance may slow down the overall development of morphological features of L2 temporality. This issue has already been discussed above and although it may appear to be simply a question of resultant slow development of L2 temporality, there may be more at play than that. The ‘slowing down’ process may also take on the form of partial or incorrect FMU mappings, which would subsequently have to be addressed in instructional settings. The third and final reason for the slow and gradual development of temporal verbal morphology is of a communicative nature. Many L2 learners may or may not be aware of morphological problems related to temporality since the first two stages of temporal expression (i.e., the pragmatic stage and the lexical stage) enable L2 learners to interact in a communicatively viable way (Bardovi-Harlig, 2000). Since communication is the ultimate goal of SLA for many L2 learners, they may not feel the need nor the motivation to produce grammatically accurate verb forms to express temporal relations. They may simply draw enough satisfaction from being able to communicate in an understandable way.

Whatever the reason(s) may be for the slow and gradual development of L2 temporality, studies have shown that instruction does have an affect on L2 learners’ developing tense-aspect systems by, for example, advancing L2 learners along the acquisitional sequence (e.g., Bardovi-Harlig & Reynolds, 1995; Cadierno, 1995; Dougherty & Varela, 1998; Harley, 1989).

Four overarching principles for the acquisition of L2 temporality have been reported in the various studies carried out as part of the European Science Foundation (ESF) project investigating the acquisition of temporality (e.g., Bhardwaj, Dietrich, & Noyau, 1988; Dietrich, 1995; Dietrich, Klein, & Noyau, 1995; Klein, 1993, 1994; Rohde, 1996). Although initially highlighted in ESF studies, the principles are so universal that they are found in other studies which investigate the acquisition of L2 verbal morphology too (e.g., Bardovi-Harlig, 1992; Bardovi-Harlig & Bofman, 1989; Kaplan, 1987). The first principle states that the development of temporal expression is both slow and gradual. L2 learners generally acquire the forms, meanings and uses of temporal features in stages and over a longer period of time. The second principle states that form often precedes function (i.e., meaning and use in our tripartite approach to grammar) (see also Salaberry, 2008). What this means in practice is that L2 learners’ verbal morphology may develop without clear meanings or uses. In other words, various forms may develop but without any clearly delineated meaning-related and/or use-related components. The third principle is that irregular verbal morphology generally precedes regular morphology. This does
not mean, however, that irregular verbal morphology is always less problematic than regular verbal morphology. It simply means that, for example, irregular past verb forms appear earlier and often include more tokens than regular past verb forms. The fourth and last principle claims that L2 learners show a tendency to avoid discontinuous tense marking (e.g., be + V-ing, have + V-en in present-day English) and initially rely on suffixed inflections before extending tense-aspect morphology to include auxiliaries.

The section above has highlighted the characteristic features of both form-oriented and meaning-oriented studies into L2 temporality in two separate discussions. So let us now turn to a comparison of both approaches.

### 3.2.4 Comparing form-oriented and meaning-oriented approaches

This section will highlight some of the main similarities and differences that may be found between the two types of approaches to L2 temporality discussed in the previous sections. In addition, the studies carried out as part of the experimental research in this dissertation (see Chapters 4, 5 and 6) will be discussed in function of the two approaches, with features of the studies being highlighted and compared in the context of the form-oriented and meaning-oriented approaches to L2 temporality.

The first similarity between both the form-oriented and the meaning-oriented approaches to L2 temporality is the view that morphology is regarded as “the surface realization of an underlying semantic system” (Bardovi-Harlig, 2000, p. 10). Although the terms form-oriented approach and meaning-oriented approach appear somewhat deceptive in that the former could be synonymous with a focus on form(s) only and the latter with a focus on meaning only, the form-oriented/meaning-oriented distinction does not reflect this difference in research focus (see Section 3.3 for more information on SLA-specific terminological issues relating to form and meaning). Both approaches take as their point of departure the assumption that temporal verb morphology is the realization of an underlying temporal system consisting of semantic concepts. The form-oriented approach investigates the underlying semantic system by means of morphology only and focuses on, for example, one specific morpheme, which is subsequently tracked in L2 learners’ interlanguages in some or all of its possible meanings and uses. In this approach, the temporal semantics are inferred from the distribution of morphemes. The meaning-oriented approach also investigates the underlying semantic system but does this by means of the three means discussed above (i.e., the pragmatic, lexical and morphological

---

The type/token distinction is one which has been taken from lexical studies, where the type-token ratio is used as a measure of lexical density. The type-token ratio may be defined as “the ratio of the total number of different words (types) to the total number of words (tokens) in a sample of text” (Crystal, 2003, p. 478). With respect to studies of verbal morphology, the type/token distinction refers to the number of different verbs (verb forms) as the number of verb types and the number of times a specific verb (form) appears as the number of verb tokens.

---

71 The type/token distinction is one which has been taken from lexical studies, where the type-token ratio is used as a measure of lexical density. The type-token ratio may be defined as “the ratio of the total number of different words (types) to the total number of words (tokens) in a sample of text” (Crystal, 2003, p. 478). With respect to studies of verbal morphology, the type/token distinction refers to the number of different verbs (verb forms) as the number of verb types and the number of times a specific verb (form) appears as the number of verb tokens.
means). Instead of focusing on, for example, one specific morpheme, the meaning-oriented approach takes as its point of departure a semantic temporal concept (e.g., ‘bygone-ness’) and investigates the means that are used—in isolation or in interaction—to express the targeted semantic concept.

The second similarity is the interlanguage perspective that both approaches take. Both the form-oriented approach and the meaning-oriented approach investigate L2 learners’ interlanguages and the development of L2 temporality in those interlanguages. In either approach, the interlanguages are described as systems which are independent of the target language(s). Although an interlanguage approach may seem at odds with the concept of instruction in studies into the acquisition of L2 temporality, there is ample room to combine the two features. Often, in instructional studies, the concept of targetlike comprehension and/or production is used as a means to assess L2 learners’ receptive and productive skills. Bardovi-Harlig (2000) addresses the relationship between target tense-aspect systems and interlanguage tense-aspect system by stating the following:

> Although in the target languages verbal morphology also encodes, with some variation, grammatical information of person and number in addition to tense and aspect, acquisition studies focus exclusively on tense and aspect. This means that an analyst will code as a good example of the preterite in Spanish . . .; in English he have studied as a good example of a present perfect. The reason for doing this is that temporal reference and person-number are distinct semantic systems. Moreover, morphology for person is acquired later than tense, at least in L2 English, . . . Thus, coding only tokens that show both tense-aspect and person-number agreement as appropriate uses of a given tense-aspect form would result in an analysis that is dependent on the emergence of person-number. (pp. 113–114)

Although Bardovi-Harlig limits her reference to the problematic tense/aspect and person/number distinctions, other issues are also at play if formal accuracy is defined as a coding/scoring criterion. As far as the experiments in this doctoral dissertation are concerned, spelling issues—such as formal regularity and irregularity—are definitely features which need to be considered. In Chapter 2, form-related problems were already addressed with respect to the past/present perfect distinction in present-day English (Section 2.4.3).

The third similarity is one which is related to the focus of investigation in SLA research. In both the form-oriented approach and the meaning-oriented approach, many—if not most—of the studies into L2 temporality have primar-

---

76 Bardovi-Harlig (2000) is judicious when making this statement and qualifies it with the following endnote: “This is not necessarily the case if a study focuses on formal accuracy in which case tense-aspect may be conflated with person and number. Unless it is explicitly stated, it is sometimes difficult to determine what scoring procedure was used. The longitudinal studies reviewed in the following section all distinguish between tense-aspect and person and number, but some of the cross-sectional studies seem to incorporate them in their focus on achieving the target.” (p. 188)
ily focused on the emergence of the simple past (or the perfective past in Romance languages). There are two main reasons for this focus on the past (Bardovi-Harlig, 2000). Firstly, the expression of pastness occurs early in the development of L2 temporality so that it can easily function as a research focus for L2 researchers and can be tracked longitudinally. Secondly, references to past situations force language learners to displace situations and such displacement generally requires some sort of marker in language learners’ primary languages and interlanguages.

The fourth similarity is related to an aspect of research design. In both the form-oriented and meaning-oriented approach, studies have generally been of a predominantly longitudinal nature. Since the studies in both approaches have investigated the emergence of L2 temporality, a longitudinal design has been clearly preferred when carrying out research. However, this does not mean that cross-sectional studies are completely absent from the two approaches. Bardovi-Harlig (2000) discusses the historical developments for a preference for studies into L2 temporality which are longitudinal in design but for the discussion at hand, these developments are not relevant.

The fifth similarity is of a conceptual nature. Bardovi-Harlig (2000) states that both form-oriented and meaning-oriented studies into L2 temporality focus on emergence, as opposed to, for example, accuracy or acquisition orders (cf. morpheme studies), which focus on acquisition. Although a valid distinction, it is not one which is consistently used in contemporary SLA research. Bialystok and Sharwood Smith (1985) and Ellis (1997) recognized two ways of operationalizing acquisition: (1) the internalization of completely new forms and (2) the increased control over forms that have already been partially acquired. However, as Ellis (2006, 2008) rightly acknowledges, acquisition may also be operationalized as “progress along a sequence of acquisition (i.e. movement from an early to later stage of development in an attested sequence)” (2006, p. 34; 2008, p. 840). Consequently, the distinction between emergence and acquisition has become somewhat blurred.

For all the similarities between form-oriented and meaning-oriented approaches into L2 temporality referred to above, there are also some fundamental differences between both approaches. The first major difference is related to the view that although morphology is regarded as the realization of underlying temporal semantic systems, it is approached differently in both approaches. Since this difference was already discussed above, no further details will be provided here.

The second difference is related to the overall focus of investigation. In this respect, Bardovi-Harlig (2000) refers to the fact that meaning-oriented studies into L2 temporality cast a broader net than studies in the form-oriented approach. This is clear when looking at Figure 3.3, which provides a side-by-side comparison of the linguistic means under investigation in both approaches. The net that is cast in the meaning-oriented approach is one which covers three
means (pragmatic, lexical, morphological), whereas the form-oriented approach focuses its attention on morphological means only and subsequently shadows various meanings and uses of one specific temporal verb form.

### CHAPTER 3

#### FORM-ORIENTED STUDIES

<table>
<thead>
<tr>
<th>What is the distribution of the emergence of verbal morphology?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological means</td>
</tr>
</tbody>
</table>

#### MEANING-ORIENTED STUDIES

<table>
<thead>
<tr>
<th>How do learners express temporal relations?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Pragmatic means</td>
</tr>
<tr>
<td>(a) scaffolding</td>
</tr>
<tr>
<td>(b) chronological order</td>
</tr>
<tr>
<td>(c) implicit ordering</td>
</tr>
<tr>
<td>(2) Lexical means: adverbials</td>
</tr>
<tr>
<td>(3) Morphological means: emergence of (verbal) morphology to mark temporal relations</td>
</tr>
</tbody>
</table>

Figure 3.3. *A comparison of form-oriented and meaning-oriented studies of temporal expression in SLA research (adapted from Bardovi-Harlig, 2000)*

Although Bardovi-Harlig (2000) clearly refers to a broader net being cast, a word of caution is warranted when interpreting her metaphor. It is definitely true that meaning-oriented studies into L2 temporality focus on various means of expressing temporality whereas form-oriented studies focus on form to the exclusion of any other means. This is probably the reason why Bardovi-Harlig uses the ‘broader net’ metaphor. However, form-oriented studies into L2 temporality partly make up for their exclusive focus on form by looking at various meanings and uses (see Figure 3.1). In any case, caution must be exercised. The ‘broader net’ metaphor should not be equated with an a priori broader research interest in L2 temporality.

The third and final major difference is related to the formulation of research questions. Because of an inherently different approach to temporal semantics, the research questions in both approaches have been different. Form-oriented studies into L2 temporality address research questions in which the formal features of L2 temporality are assumed. By contrast, meaning-oriented studies into L2 temporality do not make any assumptions with respect to such formal features. Generally, they address issues such as: How do L2 learners express (specific features of) temporality at a specific stage of development?; What are the changes in the development of L2 temporality? What are the variables that may influence the development of L2 temporality from one stage to the next?
After highlighting some of the important similarities and differences between the form-oriented and meaning-oriented approaches, it is time to investigate how the studies carried out for the research in this dissertation (see Chapters, 4, 5 and 6) fit into the body of research into L2 temporality described so far in this chapter. In other words, I would like to look at the defining features which may be used to classify the studies as either form-oriented or meaning-oriented studies. The studies that have been carried out as part of this dissertation are, in essence, examples of meaning-oriented studies, with a clear focus on two realizations (i.e., past/present perfect) of a temporal semantic notion (i.e., ‘bygone-ness’).

The target feature under investigation in all three of the studies consists of the past/present perfect distinction when used to locate bygone situations in present-day English. Bardovi-Harlig’s (2000) formulation as morphology being the “surface realization of an underlying semantic system” (p. 10) applies to all three of the studies in that I investigated the distribution of past and present-perfect verbal morphology as a means of locating bygone situations. The concept of ‘bygone-ness’ clearly delimits the semantic scope of the studies which have been carried out and places the focus of research on the investigation of one semantic concept, which is termed bygone-ness in this dissertation. Both the past and the present perfect have other meanings and uses in present-day English but the focus of investigation was solely on their meanings and uses as tenses to locate bygone situations. As already mentioned above, the past in English has, for example, secondary (i.e., extensional) meanings and uses which do not refer to pastness at all. By the same token, the present perfect also has other meanings and uses besides its meaning and use as a tense to locate bygone situations. One such example of another meaning and use of the present perfect is the continuative perfect (e.g., I have been a student for as long as I can remember), which describes a situation or habit which started in the past, continues into the present and possibly extends into the future.

In addition to focusing on the past/present perfect distinction when used to locate bygone situations in present-day English, a conscious decision was taken to focus on the meanings and uses of these two tenses and not on the forms. This is in keeping with the traditional focus of meaning-oriented studies into L2 temporality. However, this does not mean that form did not play any role. It obviously did since formal features of verbal morphology had to be used to ascertain which tenses that the participants were using. However, we consciously decided that formal accuracy was not an item of interest and were relatively lenient with formal problems that appeared during the coding and scoring procedures. Since the focus in the studies set up for this dissertation was not on formal accuracy, a decision was taken to filter out formally inaccurate answers.
on the tests and to decide on a token-by-token basis whether to disregard formally inaccurate verb forms or not.\textsuperscript{77}

All three studies were of a longitudinal nature but a note of caution is warranted at this stage. As opposed to a large number of longitudinal studies into L2 temporality, the three studies that were carried out covered a relatively short period of time with approximately three weeks between pretest and posttest sessions. The third study also contained an unannounced delayed-posttest experimental session, which took place eleven weeks after the immediate posttest. However, during that 11-week period no additional tense-related instruction was provided to any of the participants. Some scholars would point out that a three-week period between pretests and posttests may not be opportune in trying to establish the effects of instruction on the formation of temporal FMU mappings. Since many of the studies into L2 temporality investigate the emergence of specific temporal features and since the development of L2 temporality is generally considered to be slow and gradual, a longitudinal approach would appear to make more sense. However, the participants in our studies were not L2 learners whose interlanguages showed initial emergence of the past/present perfect distinction when locating bygone situations in present-day English. Many of the participants were already able to produce many correct instances of the targeted distinction, which was clear from the pretest scores. So instead of focusing on initial emergence, we were more interested in how the participants’ control over the past/present perfect distinction increased as a result of instruction. In other words, we were interested in possible signs of stabilization (with respect to the past/present perfect distinction) in the participants’ interlanguages as a result of instruction. In addition, we were not necessarily interested in completely error-free comprehension and production of the targeted past/present perfect distinction but simply in a possible improvement as a result of instruction. This approach was adopted with a view to contributing to an acquisitionally informed pedagogy. We are aware that prolonged exposure to the targeted linguistic feature may have been beneficial but within the design-related constraints such exposure was not feasible. In addition, some scholars claim that it is not necessarily the length of interaction but the intensity of the interaction which plays a vital role in establishing the effects of instruction (Klein, Dietrich, & Noyau, 1995).

3.3 Conceptualizing form-focused instruction
Throughout the introduction and Chapters 1 and 2 of this dissertation, I have made references to the terms \textit{instruction} and \textit{form-focused instruction (FFI)} without any attempts on my part to define or even operationalize the terms. However, in the experimental studies reported on in Chapters 4, 5 and 6, the reader will

\textsuperscript{77} See Chapter 4, Section 4.2.5, for additional information on the coding and scoring procedure applied in the three experimental studies carried out for this dissertation.
notice that conscious decisions were made to put into practice the types of L2 instruction provided to the various groups of participants. To be able to understand the decisions that were made, it is necessary to shed some additional light on the term form-focused instruction by pointing out some pertinent issues which SLA scholars have mulled over since, what has been referred to by Doughty and Williams (1998a) as Michael Long’s influence on the “reawakening of interest in this issue [the role of attention to form]” (p. 3). The term form-focused instruction covers a whole range of instructional setups, which cannot possibly be discussed exhaustively within the constraints of this dissertation. Consequently, I have selected those issues which I deem relevant for the discussion and research at hand. Much more elaborate discussions of FFI in general and of specific FFI-related aspects may be found in, for example, de Graaff and Housen (2009), Doughty (2003), Doughty and Long (2003), Doughty and Williams (1998c), Ellis (1994, 2002, 2006, 2008), Long (1991) and Norris and Ortega (2000).

The remainder of this second section will first consider some conceptual aspects of FFI. This will be followed by a discussion of how the effects of FFI have been addressed in the field of SLA. The discussion will be carried out using two approaches, input processing and skill acquisition theory, and will focus on the roles of input (practice) and output (practice) which proponents of both approaches espouse.

3.3.1 A conceptual framework for operationalizing form-focused instruction research

3.3.1.1 Focus on form/focus on forms

Since the renewed interest in attention to form in the field of SLA in the 1980s (see Doughty & Williams, 1998a, 1998c), the field at large has seen an enormous increase in the experimental and quasi-experimental studies into the effects of instruction that has somehow focused on form. Although this research has furthered the understanding of aspects of instructed SLA, it has also brought with it a multitude of terms which have caused confusion and, at times, disagreement.

The term form-focused instruction is used in this doctoral dissertation as an umbrella term to cover a range of possible taxonomies of pedagogical options which form instructional setups based on conceptual criteria. In this respect, I am following, among others, de Graaff and Housen (2009), who broadly define the term form-focused instruction as follows:

"[A]ny instructional activity which aims at drawing the learners’ attention to language form, where “form” stands for grammatical structures, lexical items, phonological features and even sociolinguistic and pragmatic features of language.” (p. 736)

For detailed discussions of terminological issues related to the term form-focused instruction, the reader is advised to consult, for example, de Graaff and Housen
The advantage of defining the term broadly is that it can be used multifunctionally to refer to a variety of taxonomies. Consequently, scholars are able to fall back on the term as a kind of common denominator. However, with such multifunctional use comes the problem of definitional murkiness, which was already highlighted in Chapter 2 for the concept of complexity, where it was referred to as terminological polysemy. By using such a broad term, important and necessary distinctions may be lost, resulting in vagueness when discussing fundamentally different concepts. In other words, using the term form-focused instruction may be useful but scholars should clearly introduce any further distinctions that need to be made by, for example, the specification of a more detailed taxonomy of FFI in an attempt to avoid confusion. Such specification will be undertaken in this section to prepare the reader for the experimental research reported on in Chapters 4, 5 and 6.

One such taxonomy is Long’s (1991) widely used focus on form (FonF)/focus on forms (FonFs) distinction. Although the FonFs construct has remained relatively transparent with respect to what it covers, the FonF construct has been “stretched to cover a type of FFI that it was initially intended to exclude” (Ellis, 2002, p. 15). Ellis (2002) discusses the reconceptualization of the FonF construct by comparing essential, definitional characteristics of FonF put forward by Long (1991) and definitional characteristics of FonF put forward by Doughty and Williams (1998a). Let us have a more detailed look at Ellis’s (2002) discussion of the reconceptualization of FonF. In Long’s initial definition of FonF (1991), two essential definitional features were highlighted according to Ellis (2002):

1. Attention to form takes place in the classroom where the main concern is meaning or communication.
2. Attention to form takes place incidentally as necessitated by communicative needs.

According to Ellis (2002), Doughty and Williams’s (1998b) definitional features are formulated somewhat differently, though they are not altogether incompatible with Long’s features of FonF:

1. Learners must engage with meaning before engaging with form.
2. Analysing learners’ linguistic needs is important to identify and select the forms that require instruction.
3. Instruction must be brief and unobtrusive.

When analysing Doughty and Williams’s (1998b) definitional features, it becomes clear that they highlight the advantage of a FonF approach, as opposed
to a FonFs approach or what they refer to as the “traditional forms-in-isolation type of grammar teaching” (Doughty & Williams, 1998a, p. 3), by focusing on Long’s (1991) wording “overriding focus ... on meaning or communication” (p. 46). Ellis’s (2002) concern is focused on Doughty and Williams’s second definitional feature. According to Ellis, this feature is not compatible with Long’s (1991) definitional features since the analysis of learners’ needs and the identification and selection of forms are not compatible with an incidental approach to attention to form. Ellis (2002) explains the reconceptualization from incidental to planned by highlighting researchers’ desires to conduct (quasi-)experimental studies. Such a shift, however, is of great importance to Ellis since it brings with it a different approach to treatment. Ellis (2002) states the difference as follows:

In the case of planned focus-on-form, the instruction will be intensive, in the sense that learners will have the opportunity to attend to a single, preselected form many times. In the case of incidental focus-on-form, the instruction will be extensive, because a range of linguistic forms (grammatical, lexical, phonological, pragmatic) are likely to arise as candidates for attention. ... This difference is important both theoretically and pedagogically, because it raises the question as to whether language learning benefits most from focusing on a few problematic linguistic forms intensively or from a scatter-gun approach, where multitudinous problematic forms are treated randomly and cursorily and where treatment may or may not be repeated. (p. 16, quoted with original highlighting)

Consequently, Ellis (2002) suggests that FFI should be conceptualized more appropriately in terms of a more nuanced three-way distinction (i.e., FonFs, planned FonF, incidental FonF) and not in terms of a crude two-way distinction (i.e., FonFs, FonF). Table 3.1 provides an overview of Ellis’s three-way conceptualization of FFI.

<table>
<thead>
<tr>
<th>Type of FFI</th>
<th>Primary focus</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Focus-on-forms</td>
<td>Form</td>
<td>Intensive</td>
</tr>
<tr>
<td>2. Planned focus-on-form</td>
<td>Meaning</td>
<td>Intensive</td>
</tr>
<tr>
<td>3. Incidental focus-on-form</td>
<td>Meaning</td>
<td>Extensive</td>
</tr>
</tbody>
</table>

Although Ellis’s three-way conceptualization of FFI is a valid one, it does open the door to a possible free-for-all with respect to categorizing specific types of FFI if concepts such as meaning and communication are stretched. The planned-FonF category is a category which could possibly cause some confusion in that it allows the preselection of a linguistic form (just like the FonFs category). However, this does not mean that planned, contextualized instruc-
tional tasks are automatically of a planned-FonF nature. The instructional practice-based setups in Study 1 (Chapter 4), Study 2 (Chapter 5) and Study 3 (Chapter 6) are without a doubt setups which were planned and contextualized, but using Ellis’s taxonomy it could be argued—according to Ellis’s approach—that they are not pedagogical options for planned-FonF instruction since the context in which they are presented is not necessarily a communicative one. Indeed, this is in line with one of Long’s (1991) initial definitional characteristics, which refers to the required presence of a meaning or communication component when focusing on form. Using Ellis’s (2002) conceptualizations of FFI, the instructional setups in Chapters 4, 5 and 6 could be referred to as functional or situational since the instructional materials provide L2 learners with the opportunity to practice producing specific target features in a situational context. Although such activities are functional, they are not inherently communicative. Consequently, such materials would fall under FonFs instruction in Ellis’s taxonomy because the main concern is on form rather than meaning despite the apparent focus on meaning (Ellis, 2002). In addition, learners who are confronted with functional tasks are fully aware that the target of instruction is accurate use through practice (i.e., repeated use) of a specific target feature (Ellis, 2002). In other words, it could be stated that the participants in all three of the studies conducted for this doctoral dissertation were not pushed to communicate but rather to interact with instructional tasks of a non-communicative nature. In addition, the participants knew in advance what the focus of instruction was by participating in a theoretical experimental session, which preceded any practice-based instruction. The line between communicative context and functional/situational context is a narrow one and may possibly create confusion for many. However, what must not be forgotten is that for FFI instruction to be categorized as planned FonF, preselection must still entail a communicative component. Focused communicative tasks, for example, are tasks that have been selected to elicit the production of a specific target feature in the context of performing a communicative task. Ellis (2002) clearly states that such tasks have “all the characteristics of communicative tasks” (p. 21). If one looks in detail at the nature of the instruction and the practice-based instructional components in Studies 1, 2 and 3, it is clear that the majority of the participants were somehow pushed to communicate by means of picture selection tasks. Indeed, these tasks were available for both input-practice and output-practice groups. The only group of participants which was not pushed to communicate was the input-only group in Study 1. As a consequence of this design-related feature, the pedagogical tasks for the tasks used for the practice-based training in the three studies reported on in this doctoral dissertation may be categorized as explicit FonF tasks using Ellis’s nuanced taxonomy.
3.3.1.2 Questions raised in L2 type-of-instruction research

Although the focus of contemporary L2 instruction research appears to be firmly focused on what types of instruction are most effective for learning L2s, this was not always the main focus of investigation in the past. In the 1980s and early 1990s, much of the L2 instruction research was focused on whether or not instruction made any difference at all. Consequently, many studies investigated the effects of instruction by comparing instructed L2 learners with un instructed L2 learners. For the purposes of this doctoral dissertation, I will focus on features of L2 type-of-instruction research and take for granted the turn in research focus with respect to the value of L2 instruction in comparison with naturalistic L2 learners.

In their oft-cited research synthesis and meta-analysis with regard to the effectiveness of L2 instruction, Norris and Ortega (2000, pp. 418–419) provide the following six general research questions which, according to them, L2 type-of-instruction research has sought to investigate:

1. Is an implicit or an explicit approach more effective for short-term L2 instruction?
2. Can raising learners’ metalinguistic awareness of specific L2 forms facilitate acquisition by fostering psycholinguistic processes of form-to-function mapping?
3. Is instruction that draws learners’ attention to relevant forms in the context of meaning-focused lessons more effective than an exclusive focus on meaning and content?
4. Is negative feedback beneficial for L2 development, and if so, what types of feedback may be most effective?
5. Is acquisition promoted more effectively when learners process the input in psycholinguistically relevant ways than when they experience traditional grammar explanation and practice?
6. Is comprehension practice as effective as production practice for learning L2 structures?

A discerning reader will notice that the six general research questions do not represent disparate research interests. Norris and Ortega (2000) point out the common (cognitive) thread throughout the research questions by stating that type-of-instruction studies addressing them are based on a single theoretical assumption, that is, the aim of instructional interventions to bring about changes in how L2 learners focus their attention when processing L2s to facilitate noticing of linguistic features and, subsequently, acquisition. In addition, Norris and Ortega (2000) state that these changes should be brought about “in efficient ways in terms of rate of acquisition and target-like levels of ultimate attainment” (p. 419). Consequently, one of the major concerns of contemporary L2 type-of-instruction research is centred on the importance of implicit
and/or explicit cognitive processing involved in optimal L2 learning (Norris & Ortega, 2000).

### 3.3.2 The effects of practice-based instruction according to input processing and skill acquisition theory

The discussion in this section will focus on two issues: (1) the essential components of two common approaches available in contemporary SLA research, which are referred to using the terms *input processing* and *skill acquisition theory*, and (2) the instructional realizations of these approaches in what Norris and Ortega (2000) have referred to as “practice-based instructional types” (p. 422). As such, this section is not meant to serve as an exhaustive discussion of the approaches. Within the constraints of this dissertation justice cannot possibly be done to every detail of the approaches. I will present the issues which are relevant for the research carried out and I will highlight—wherever necessary—any useful references for additional information about the approaches. After discussing the essential components of both approaches, I will place the focus on the role of practice in general and on the roles of both input (practice) and output (practice) in both approaches. It is important to underline that the input/output distinction should not be interpreted as an either/or distinction. Both input and output are features of both approaches. However, because the approaches have different views on the roles of both, a discussion of how the roles of input (practice) and output (practice) are viewed in both approaches is warranted. The order in which the approaches are discussed does not reflect any preference on my part for one of the two approaches.

#### 3.3.2.1 Input processing and processing instruction

**Input processing.** The name most associated in SLA research with input processing (IP) and its pedagogical realization, which has been termed *processing instruction* (PI), is Bill VanPatten. The IP and PI research paradigms were launched in the early 1990s and the article which first introduced the concepts is VanPatten and Cadierno (1993). Both concepts have established themselves in contemporary SLA research with a relatively large body of research investigating both general and specific features of IP and PI (e.g., Benati, 2004; Benati & Lee, 2008; Cadierno, 1995; Farley, 2001a, 2001b, 2004a, 2004b; Lee & Benati, 2008).

---

*The debate about cognitive processing is reflected in the interface debate in SLA, which addresses the interfaces between explicit and implicit knowledge. Three positions are generally distinguished in this debate: (1) the no-interface position (e.g., Hulstijn, 2002; Krashen, 1981, 1985, 1994; Krashen & Terrell, 1983; Paradis, 2004), (2) the weak-interface position (e.g., Ellis, 1990, 1997, 2008) and (3) the strong-interface position (e.g., DeKeyser, 1997, 1997, 2001, 2003; DeKeyser & Jufí, 2005; Sharwood Smith, 1988). For the discussion at hand, some form of interface is required since practice would otherwise play no role—or only an extremely small role—in the acquisition of L2 features. The importance awarded to the independent variable practice in the three studies in Chapters 4, 5 and 6 reflects the assumption that some form of interface exists.*
At the same time, questions about IP and PI have also been raised. In turn, these questions have not only spawned interesting debates between SLA scholars; they have also provided insightful remarks about the theoretical and practical foundations of IP and PI (e.g., DeKeyser, Salaberry, Robinson, & Harrington, 2002; Doughty, 2004; Lightbown, 2004; VanPatten, 2002b, 2002c). The descriptions provided in this section are meant as summaries of the general features of both IP and PI within the context of the research being carried out for this dissertation.

In the IP research paradigm, language acquisition has generally been described as “the development of some underlying competence on which skills in language use depend” (VanPatten, 2004a, p. 29), with VanPatten (1996) clearly defining the developing system as “the complex of mental representations that as an aggregate constitutes the learner’s underlying knowledge of the second language (phonology, syntax, morphology, etc.)” (p. 9). In addition to defining the developing system, VanPatten has also referenced to other names for the developing system in SLA such as underlying mental representation, interlanguage or developing system. Like many other SLA scholars, VanPatten is extremely consistent in stressing that second language acquisition is inherently complex. One of the several causes of this complexity is the fact that acquisition does not consist of one single process but rather of multiple processes, multiple knowledge domains and multiple interactions of both, at all possible levels of acquisition (VanPatten, 2004b, 2007). In other words, acquisition is said to be a multicomponential phenomenon. The three major sets of SLA processes that are generally distinguished in the IP research paradigm are: (1) input processing, (2) accommodation and (3) restructuring, which are represented schematically in Figure 3.4 (VanPatten, 2003, 2004b, 2007). Figure 3.4 has been taken from more recent IP-related publications. The reader will notice that the representation is actually somewhat misleading. Numbers 1, 2 and 3 in Figure 3.4 do not actually refer to the aforementioned three major sets of SLA processes (i.e., input processing, accommodation, restructuring). Number 2 in Figure 3.4 combines the second and third sets of processes (i.e., accommodation, restructuring) and number 3 refers to a process—which in turn consists of a set of subprocesses—which results in output. However, as far as VanPatten and IP/PI proponents are concerned, output is not directly one of the SLA processes. This does not mean, however, that output is sidelined in the IP model. It is given a role which will be highlighted below. A clearer schematic representation would be either a simplified version as found in the earlier IP-related publications (e.g., VanPatten, 1996) or an adapted representation as found in Figure 3.5.
The focus of VanPatten’s and IP/PI proponents’ discussions has been on one of the three major processes, namely, on input processing. Although VanPatten has provided information—albeit relatively limited information—on the processes of accommodation and restructuring, the focal point of his and IP/PI advocates’ publications has been on input processing. This may appear to be a rather trivial fact but it is important to underscore at this point. VanPatten himself stresses this by regularly pointing out that his model or theory of input processing should not actually be equated with a model or theory of SLA since SLA implies more than simply processing input (VanPatten, 2004a, 2004b, 2007). Central to VanPatten’s model of input processing (and, by extension, to model-based pedagogical proposals) are “assumptions about the nature of “attention,” language “processing” and the structure of attentional and memory “resources” (DeKeyser et al., 2002, p. 806).

Accommodation is described by VanPatten (2004a) as “either the partial penciling in or complete incorporation of a surface feature (form-meaning connection) of language into the developing system” (p. 33) and restructuring as “what may happen to the developing system after a form has been accommodated” (p. 33). Whereas accommodation involves quantitative changes to the developing system or linguistic behaviour, restructuring involves qualitative changes to the developing system of linguistic behaviour (VanPatten, 2004a).
As is clear from Figure 3.4, IP is the process which focuses on how input is processed and how a subset of the input, referred to as intake, is subsequently made available for further processing in SLA. As mentioned above, the three major processes may themselves contain subprocesses. In VanPatten’s model, IP is said to consist of two subprocesses, which are not necessarily mutually exclusive: (1) the formation of form-meaning connections (FMCs) and (2) parsing (VanPatten, 2004a). The formation of FMCs occurs the very first time that an L2 learner makes a connection between form and meaning. In other words, it occurs the first time that an L2 learner maps form onto meaning or vice versa. The nature of the connection that is made may vary from whole to partial, from correct to incorrect (VanPatten, 2004a, 2004b). Although VanPatten provides ample data from, for example, Spanish to highlight partial connections, I will use data from English in the context of this dissertation. Imagine that a Dutch-speaking ESL learner comes across the verbal form is diving and processes the relationship as an FMC between the English verb inflection be (present) + -ing (form) and the present (meaning). This connection is, of course, correct but it is partial or incomplete since the verb inflection be (present) + -ing in present-day English encapsulates not only a temporal meaning (i.e., present) but also an aspectual meaning (i.e., progressive). Subsequent processing would then be required to complete the FMC and to integrate the aspectual meaning into the already established FMC with a temporal meaning.

The processing of input is not a random process and VanPatten has developed a series of processing principles in an attempt to describe this complex feature. A complete discussion of all of the principles falls outside the scope of this dissertation but for an overview of one of the main principles and some of its subprinciples, the reader is advised to read Section 2.4.5 in Chapter 2. The second process found in input processing is referred to as parsing. According to VanPatten (2004a) parsing refers to “how learners assign syntactic categories to words they comprehend and to what kind of syntactic representations learners build during comprehension” (p. 31). In his discussions of parsing, VanPatten (1996, 2003) generally refers to the first-noun strategy, which English learners of Spanish rely on when processing Spanish sentences, which in 60 per cent of Spanish sentences is unproblematic since SVO and even VO—Spanish is a pro-drop language—are interpretable by an English L1 parser. However, the remaining 40 per cent of Spanish sentences follow other syntactic patterns (e.g., SOV, OVS, OV), resulting in regular misinterpretations of OVS sentences as SVO sentences.

Although VanPatten refers fairly regularly to form-meaning connections (FMCs), he does occasionally highlight the function aspect of connections (or mappings) too (e.g., VanPatten, 2004a, 2004b). In light of the tripartite approach to mappings adopted in this doctoral dissertation, the reader is advised to read VanPatten’s form-meaning connection (FMC) as form–meaning–use mapping (FMU mapping), which is the term used in this dissertation.
VanPatten’s model of IP is essentially a limited-capacity model of IP, which forces L2 learners to attend to input data. Attending to input data takes place—following VanPatten—according to specific principles, which have been formulated as the basis of VanPatten’s model of IP. By way of summary, a 2002 version of VanPatten’s main principles of IP is provided in Table 3.1.

### Table 3.2. Main principles of input processing (from VanPatten, 2002b, p. 758)

<table>
<thead>
<tr>
<th>P1.</th>
<th>Learners process input for meaning before they process it for form.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1a.</td>
<td>Learners process content words in the input before anything else.</td>
</tr>
<tr>
<td>P1b.</td>
<td>Learners prefer processing lexical items to grammatical items (e.g., morphology) for the same semantic information.</td>
</tr>
<tr>
<td>P1c.</td>
<td>Learners prefer processing “more meaningful” morphology before “less” or “nonmeaningful” morphology.</td>
</tr>
</tbody>
</table>

| P2. | For learners to process form that is not meaningful, they must be able to process informational or communicative content at no (or little) cost to attention. |

<table>
<thead>
<tr>
<th>P3.</th>
<th>Learners process a default strategy that assigns the role of agent (or subject) to the first noun (phrase) they encounter in a sentence/utterance. This is called the first-noun strategy.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3a.</td>
<td>The first-noun strategy may be overridden by lexical semantics and event probabilities.</td>
</tr>
<tr>
<td>P3b.</td>
<td>Learners will adopt other processing strategies for grammatical role assignment only after their developing system has incorporated other cues (e.g., case marking, acoustic stress).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P4.</th>
<th>Learners process elements in sentence/utterance initial position first.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P4a.</td>
<td>Learners process elements in final position before elements in medial position.</td>
</tr>
</tbody>
</table>

### Input (practice) and output (practice). Let us now turn to the roles of input (practice) and output (practice) in IP before looking at the IP-based pedagogical realization, which is also known as processing instruction (PI). In VanPatten’s multicomponential view of SLA, both input (practice) and output (practice) are featured. However, the roles that both concepts play in SLA are fundamentally different and are described by VanPatten (2004b) as “complementary” (p. 6). In fact, according to VanPatten, SLA is input-dependent but a form of dependency cannot be ascertained for the relationship between SLA and output. VanPatten (VanPatten, 2004a) discusses this issue by stating that input is at some level the “primary initial ingredient for the development of competence, however one construes that competence” (p. 35). Indeed, the role of input is universally accepted in SLA and in SLA theories (see Gass, 1997, and for a contemporary systematic discussion of the role of input across SLA theories,
As far as the role of output is concerned, VanPatten (2003, 2004a, 2004b) is not an advocate of claims that SLA is in some way output-dependent. The adjective *output-dependent* implies that without any output, there would be no SLA, or at least, no successful SLA. However, according to VanPatten's views, output is not an essential part of the acquisition process since it does not bring about any changes to the developing system. Using Salaberry's (1997) words with respect to VanPatten and Cadierno's (1993) claim about the acquisition/learning distinction: “Acquisition and learning are the accurate description of the cognitive processes generated by input and output practice” (p. 423). This confirms the essential, first-class role of input (practice) and the non-essential, second-class role of output (practice) in second language acquisition.

In his discussion of the roles of output, VanPatten (2004a, 2004b) distinguishes two types of output: (1) output as interaction and (2) output as production (or Swain's (1985) notion of pushed output). Output as interaction is output which is used in interactional contexts. What this type of output does, according to VanPatten (1996, 2004a, 2004b), is change—in interactional contexts—the so-called task demands which are placed on L2 learners during input processing, resulting in freed up attentional resources, which, in turn, give L2 learners the opportunity to process input elements which they may have missed initially. In other words, input is made more manageable and quoting VanPatten (2004b): “Greater manageability can lead to increased resources for noticing” (p. 26). This role of output is, of course, based on VanPatten’s view of input processing being a limited-capacity phenomenon. As a result of interacting with an interlocutor, the L2 learner is given essential data—which are found in the interlocutor’s output—which function as input for the L2 learner. In addition, output as interaction may also lead L2 learners to noticing that something which they have said is not the same as what they have heard during the interaction, resulting in the noticing, which is essential for making FMCs. If we turn to output as production, we see that VanPatten is once again extremely clear about his point of view. He disagrees with the assumption that using a form in one’s output is a direct path to second language acquisition (VanPatten, 2003, 2004a, 2004b). VanPatten (2003, 2004a, 2004b) simply acknowledges that output may be beneficial in developing language skills (e.g., accuracy and fluency) but skill development and the creation of an implicit linguistic system are two separate phenomena.

**Processing instruction (PI).** In a 2005 publication on PI, VanPatten addresses the PI background by linking the pedagogical intervention to his general model of input processing. The fundamental question that PI attempts to address is—using VanPatten’s (2005) words—“Is there a way to enrich learners’ intake using insights from input processing?” (p. 272) or a reformulation of this question: “To what degree can we either manipulate learner attention during input processing or manipulate input data so that more and better form-meaning connections are made?” (p. 272). Generally, PI is discussed in function
of its three basic characteristics, which VanPatten (2005, p. 273) describes as follows:

1. Learners are given information about a linguistic structure or form.
2. Learners are informed about a particular IP strategy that may negatively affect their picking up of the form of structure during comprehension.
3. Learners are pushed to process the form or structure during activities with structured input — input that is manipulated in particular ways so that learners become dependent on form and structure to get meaning (i.e., learners are pulled away from their natural processing tendencies toward more optimal tendencies).

VanPatten and other IP/PI advocates (e.g., Benati & Lee, 2008; Lee & Benati, 2007; VanPatten, 1996; Wong, 2004b) have described PI using alternative terminology but the characteristics described above are consistently recognizable in the various descriptions of PI. Some descriptions have used abbreviations, which have become common in discussions of PI. For example, the information about the linguistic structure or form has been termed explicit information (EI). The activities with structured input have been termed structured input activities (SIAs). What is clear from the characteristics above is that PI is essentially an input-based approach to (grammar) instruction, in which input and input practice play pivotal roles. The way in which PI fundamentally differs from other input-based treatments is found in the second PI characteristic: the identification of a possibly problematic processing strategy based on the model of input processing highlighted above. VanPatten (2005) encapsulates this feature by stating that “PI does not just determine what is a problem form or structure but also why it is a problem vis-à-vis the processing issues” (p. 275). In addition to explaining the nature of the three PI components, VanPatten and PI advocates often discuss, in detail, the guidelines that have been drawn up for developing SIAs. Wong (2004b) describes the development of SIAs in function of two general procedures that must be followed: (1) identify the processing problem/strategy and (2) follow the guidelines for the development of SIAs. In one of the earliest publications on PI, VanPatten (1996) provides the following six guidelines for drawing up SIAs:

---

81 Within the PI paradigm, VanPatten (1996, 2005) refers to two types of SIAs: (1) referential SIAs and (2) affective SIAs. Referential SIAs are activities for which right or wrong answers exist and for which L2 learners must rely on targeted forms to get meaning. Affective SIAs are activities for which L2 learners are invited to express opinions, beliefs or to give other affective responses. In affective SIAs, L2 learners are involved in processing information about the real world.
1. Present or teach only one thing at a time.
2. Keep meaning in focus.
3. Have L2 learners do something with the input.
4. Use both written and oral input.
5. Move from sentences to connected discourse.
6. Keep the psycholinguistic processing strategies in mind.

Although VanPatten (1996) himself has clearly stated that the guidelines above are to be viewed as guidelines and not maxims and that the application of the guidelines may vary from lesson to lesson, their realization in SIAs has become somewhat problematic. Doughty (2004) highlights considerable changes between SIAs in earlier and later studies into PI, which she describes as “the departure from the original PI treatment design” (p. 262). The biggest change—according to Doughty (2004)—is the move from processing activities (in earlier studies) to language manipulation and metalinguistic activities (in later studies), which has led to a loss of focus on meaning and function during learner processing of the target feature. Doughty (2004) puts forward two possible reasons for this departure by focusing on the actual processing problems that have been addressed in later studies into PI. Whereas initial studies focused on what has been referred to by VanPatten as the first-noun strategy, later studies have focused on two types of processing problem: (1) L2 learners processing input using strategies developed during L1 acquisition (e.g., first-noun strategy) and (2) L2 learners failing to notice features that are not salient in the input, either because they are too difficult or because the information is available elsewhere in more noticeable forms, or both. The consequence has been that the SIAs in later IP studies have reflected more directly a FonFs approach whereas the SIAs in earlier IP studies were developed with a clearer FonF approach. In turn, Doughty (2004) claims that the emphasis on FonFs in later PI studies has resulted in gained knowledge which is not truly relevant to real SLA processes.

The IP and PI research paradigms have resulted in a plethora of publications. However, for what VanPatten (2005) has referred to “robust results” (p. 276), issues of contention related to the paradigms have also been discussed, resulting in warranted notable observations by various SLA scholars on, for example, theoretical and definitional issues, on design and operationalization issues, and on replication issues (e.g., DeKeyser et al., 2002; Doughty, 2004; Lightbown, 2004; Salaberry, 1997).

3.3.2.2 Skill acquisition theory

Skill acquisition theory principles. In their 2010 publication *Key Terms in Second Language Acquisition*, VanPatten and Benati refer to Robert DeKeyser as “the name most associated with skill and skill learning in SLA” (p. 150). Skill acquisition theory in SLA—like many other cognitive approaches to SLA—does not originate from the field of SLA itself but rather from the field of cog-
nitive psychology, where its study has been and still is substantial (e.g., Anderson, 1983, 1993, 2000; Carlson, 1997, 2003; Newell & Rosenbloom, 1981). Since the mid-1980s, skill acquisition theory has featured in SLA and has been productive in guiding SLA research (Ortega, 2009). Because of the close link with endeavours in cognitive psychology, a description of skill acquisition theory in SLA irrevocably falls back on concepts and issues which are also found in cognitive psychology. At the same time, however, the wholesale importation of skill acquisition theory and research interests into the field of SLA is not recommended because of the language-specific features and research interests that come with learning an L2 and language skills.

DeKeyser (2007c) describes the basic claim of skill acquisition theory as follows:

The basic claim of skill acquisition theory is that the learning of a wide variety of skills shows a remarkable similarity in development from initial representation of knowledge through initial changes in behavior to eventual fluent, spontaneous, largely effortless, and highly skilled behavior, and that this set of phenomena can be accounted for by a set of principles common to the acquisition of all skills. (p. 97)

In essence, skill acquisition theory is an information processing theory which describes the gradual transformation of performance from controlled to automatic (Ortega, 2009). However, the transformation is not of a haphazard nature. It is predicated on specific claims and constructs, which interact to form the principles of skill acquisition theory. The transformation essentially takes place as a result of practice over a series of trials. In turn, such practice ensures that the number of controlled processes is diminished whilst simultaneously increasing the number of automatic processes during performance (Ortega, 2009).

Three developmental stages in the skill-acquisition continuum have generally been recognized and have been referred to using different taxonomies. The most common terminology in SLA research for the three developmental stages has been (1) declarative, (2) procedural and (3) automatic (taken from Anderson’s earlier formulations of his Adaptive Control of Thought (ACT) theory (Anderson, 1983)).

Figure 3.6 provides a schematic overview of the

---

82 In this respect, skill acquisition theory in SLA is no different from VanPatten’s model of input processing in SLA, which also borrows from, for example, the field of cognition.
83 DeKeyser (2007a) underscores the presence of other taxonomies in discussions of skill acquisition and refers, for example, to the following three taxonomies: (1) cognitive, associative and autonomous, (2) declarative, procedural and automatic and (3) presentation, practice and production.
84 Anderson’s ACT model is a complex model of skill acquisition which has been updated on several occasions. Within the constraints of this dissertation relevant aspects have had to be selected for discussion. This has happened at the expense of other features of the model. For much more extensive discussions of the ACT model, including its various versions (e.g., ACT*,...
developmental stages in skill acquisition theory. The difference between the three developmental stages may be found in both the type of available and/or used knowledge and the ways in which that knowledge is used.

\[
\begin{array}{ccc}
\text{declarative knowledge} & \rightarrow & \text{procedural knowledge} & \rightarrow & \text{automatic knowledge} \\
1 = & \text{proceduralization} & 2 = & \text{automatization}
\end{array}
\]

Figure 3.6. Developmental stages in skill acquisition theory

The three developmental stages in Figure 3.6 represent the fundamental ideas behind general skill learning, which show a striking resemblance to the mastery of L2 skills (Dörnyei, 2009). To understand the nature of skill learning, let us have a closer look at these three developmental stages. The declarative (or cognitive) stage is the stage which is characterized by declarative knowledge (‘knowledge that’). In essence, such declarative knowledge is knowledge which has not yet been acted on (DeKeyser, 2007a, 2007c). What the learner is required to do is undertake the initial encoding of the skill by acting on the declarative knowledge that is provided. In, for example, an instructed SLA context, the declarative stage would be the stage at which L2 learners are given explicit information about an L2 target feature. In the second stage, the procedural (or associative) stage, learners are invited to practise the encoded skill. Progression is characterized by a shift from relying on declarative knowledge to procedural knowledge. In an instructed SLA context, L2 learners are invited at this stage to act on declarative knowledge by turning ‘knowledge that’ into ‘knowledge how’. In other words, L2 learners are pushed to turn declarative knowledge into procedural knowledge. Anderson (1976, p. 117) described the differences between declarative and procedural knowledge by means of three assumptions:

1. Declarative knowledge seems to be possessed in an all-or-none manner, whereas procedural knowledge seems to be something that can be partially possessed.
2. One acquires declarative knowledge suddenly, by being told, whereas one acquires procedural knowledge gradually, by performing the skill.

3. One can communicate one’s declarative knowledge verbally, but not one’s procedural knowledge.

DeKeyser (1997) has found that the process of proceduralization, that is, transitioning from declarative knowledge to procedural knowledge, need not take up too much time. In other words, proceduralization may be complete after a few trials or instances. Of course, proceduralization can be effectuated only if certain conditions are available (DeKeyser, 2007c). For example, the relevant declarative knowledge should be made available and should be drawn on in attempts to execute any target behaviour. In addition, the proceduralization of declarative knowledge also shows a great deal of regularity, which has been captured in the central concept referred to as the power law of practice or power law of learning. This concept refers to the regularity which captures the decrease of reaction time and error rate as a consequence of practice.\(^{85}\)

The huge advantage of procedural knowledge over declarative knowledge, according to DeKeyser (2007c), is that procedural knowledge “no longer requires the individual to retrieve bits and pieces of information from memory to assemble them into a “program” for a specific behavior; instead, that program is now available as a ready-made chunk to be called up in its entirety each time the conditions for that behavior are met” (p. 98). What follows the procedural stage is a generally long road to automatic (or autonomous) knowledge, which represents the third stage of skill learning. This road paves the way for increased robustness and fine-tuning of the procedural knowledge so that that knowledge may be used for behaviour that shows high levels of consistency, fluency and spontaneity, which is termed automaticity in skill acquisition theory terminology. When applying the declarative/procedural/automatic taxonomy to language learning, Anderson saw the differences between L1 and L2 leaning as simply a difference in the stage that had been reached (Ellis, 2008). Whereas L1 learners almost always progress to the automatic stage, L2 learners often reach only the procedural stage. Or as Ellis (2008) puts it: “[A]lthough foreign language learners achieve a fair degree of proceduralization through practice, and can use L2 rules without awareness, they do not reach full autonomy” (p. 429).

Although skill learning theory appears relatively straightforward, it is a theory which explores a highly complex phenomenon. In turn, this phenomenon is influenced by a variety of factors. In addition to the inherent complexity of skill learning, there are also competing views in SLA research with respect to the fundamental interaction between declarative and procedural knowledge and the ultimate development of procedural knowledge, which has also been referred to as the declarative-to-procedural shift (Dörnyei, 2009). The three best-known

---

\(^{85}\) See, for example, DeKeyser (2001, 2007a, 2007c) for additional information about the power law of practice in an SLA context.
views on the development of procedural knowledge are: (1) the proceduraliza-
tion of declarative knowledge, which was described above (e.g., Anderson,
1983, 1993, 2000), (2) the construction of parallel procedural knowledge, a view
which maintains that declarative and procedural knowledge are independent,
resulting in an inability to convert declarative knowledge into procedural. (e.g.,
Hulstijn, 2002; Ullman, 2005) and (3) the construction of a repertoire of epis-
odic instances (e.g., Logan, 1988, 2005). Because of the constraints of this
dissertation, no further details about this highly interesting yet complex dis-
cussion will be provided. For detailed overviews and discussions, the reader is
advised to consult, for example, DeKeyser, 2001, and Dörnyei, 2009.

Input (practice) and output (practice). At the basis of the gradual trans-
formation from declarative knowledge to procedural knowledge, and from
procedural knowledge to automatic knowledge, lies the concept of practice in
skill acquisition theory. DeKeyser (2007a, 2007c) highlights the concept of
practice in SLA by discussing its role and possible features in skill acquisition
theory. The role that is assigned to practice in skill acquisition theory is pivotal
since practice is the driving force behind moving L2 learners through the vari-
cious types of knowledge on their way to complete L2-learner autonomy. Unlike
VanPatten’s model of input processing, which views SLA as input-dependent
and assigns no role to output (practice) in the acquisition process,86 skill acquisi-
tion theory takes a different approach to practice. It regards practice in general
as a necessary component of the SLA process but underscores that the gradual
transformation, which is called proceduralization, is skill-specific (Ortega, 2009).
DeKeyser (2007a) acknowledges that skill specificity is probably the issue which
has attracted most of the research focus in applied linguistics lately. Following
skill acquisition theory tenets, only minimal or no transfer would take place
between skills. In the context of practice in instructed SLA, this would lead to
skill-specific practice effects for both input practice and output practice. In
other words, input practice would lead to the acquisition of comprehension-
based (or receptive) skills whereas output practice would lead to the acquisition
of production-based (or productive) skills.87 This, of course, is not in agreement
with VanPatten’s predictions about practice, which state that only input (prac-
tice) is necessary for acquisition to take place, that is, the acquisition of recep-

86 VanPatten’s model of input processing takes output (practice) into consideration but only as a
measure to improve accuracy and fluency (output as production) or as a facilitative feature (out-
put as interaction) and not as a direct route to SLA.
87 Research in cognitive psychology has also addressed the issue of transfer appropriate process-
ing (TAP), which refers to a key concept underlying the nature of encoding/learning-retrieval
relationships (Segalowitz, 2010). The approach to practice in skill learning theory described above
fits into the broader picture of learning language skills and TAP. Lightbown (2008) describes the
underlying tenet of TAP as follows: “The fundamental tenet of TAP is that we can better re-
member what we have learned if the cognitive processes that are active during learning are similar
to those that are active during retrieval” (p. 27). For a more detailed discussion of the TAP in
instructed SLA, see Lightbown, 2008, and, Segalowitz, 2010.
tive and productive skills. Output (practice) simply functions to increase accuracy and fluency or has a facilitative role.

3.4 Conclusion
The overall focus of this chapter was on the grammatical category of tense and the ways in which it interacts with the field of instructed SLA. In the first section (Section 3.2), I provided the background for how tense has been conceptualized in studies investigating the acquisition of L2 temporality. After presenting—in isolation and by means of comparison—the two approaches (i.e., focus-oriented approach and meaning-oriented approach) that have been used to investigate L2 temporality, I focused on the meaning-oriented approach, which was used for the experimental studies that were carried out for this dissertation. Although studies with a meaning-oriented approach show specific design-related and methodological features, they do not specifically shed any light on possible instructional setups that may be drawn up for investigation.

In an attempt to highlight relevant research-related issues to the instruction that was provided to the participants in the experimental research, I turned the focus to instructed SLA in the second section of this chapter (Section 3.3). The term form-focused instruction (FFI) was explained and with it possibilities for conceptualizing this umbrella term. Since the instruction in Studies 1, 2 and 3 was of an FFI nature, it was important to contextualize this term and to provide a backdrop for the experimental research carried out. Subsequently, I turned to the concept of practice, which was used as a way of distinguishing between the major instructional setups in the studies that were carried out. Since the distinction between input practice and output practice formed the foundation of the experimental research that was carried out in this dissertation, I tried to indicate the roles and the importance of practice by discussing two approaches—input processing and skill acquisition theory—which both assign crucial roles to the concept of practice but do so in fundamentally different ways. Input processing underscores the importance of input (practice) in SLA and the complementary role of output (practice) in increasing accuracy and fluency and facilitating the production of input. Consequently, SLA is said to be only input-dependent according to input processing tenets. Skill acquisition theory, on the other hand, embraces both input (practice) and output (practice) as essential components in SLA. Both forms of practice lead to the gradual transformation of performance from declarative knowledge to procedural and from procedural to automatic, that is, from controlled to automatic. Input processing and skill acquisition theory share some cognitive features (e.g., importance of practice, provision of explicit L2 information (declarative knowledge) before attempting to continue with proceduralization) but skill specificity is not one of them. Whereas input processing does not advocate skill specificity and claims that transfer from input-practice effects does exist and that both comprehension and production in
SLA are influenced by input practice, skill acquisition theory is more reluctant and promotes the idea that practice effects are skill-specific. In other words, according to the principles of skill acquisition theory, input practice leads to the acquisition of comprehension-based (or receptive) skills and output practice to the acquisition of production-based (or productive) skills. As a result of this fundamental difference with respect to skill specificity, hypotheses for the experimental research carried out are different depending on which approach is taken.

The following three chapters (Chapter 4, 5 and 6) will report on the three studies that were carried out for this doctoral dissertation. In addition to providing design-related and methodological information for every study, the reader will be guided through the statistical analyses and preliminary discussions of the results obtained. More detailed discussions and interpretations will be provided in Chapter 7, which will also address the strengths and limitations of the studies and will shed light on pedagogical implications and options for possible future research.
CHAPTER 4
STUDY 1

The great tragedy of science – the slaying of a beautiful hypothesis with an ugly fact.
(Thomas H. Huxley)

4.1 Introduction
After reviewing relevant issues with respect to the complex ESL target structures under investigation in this doctoral dissertation, that is, the past/present perfect distinction when locating bygone situations in present-day English, I will now place the focus on the experimental nature of the research carried out. This chapter is the first of three chapters dedicated to experimental studies into the effects of explicit form-focused instruction (FFI). The general research question in all three of the studies was whether explicit FFI had any effect on the acquisition of the aforementioned complex past/present perfect distinction by Dutch-speaking ESL learners. In addition, the studies sought to investigate the differential effects of various forms of explicit FFI on the acquisition of the past/present perfect distinction. For Study 1, the following three research questions were formulated:

Research questions
1. Does explicit form-focused instruction (FFI) have an effect on (Dutch-speaking) ESL learners’ performance with regard to complex temporal FMU mappings (i.e., the past/present perfect distinction) in L2 English?
2. If explicit FFI does have an effect, can any overall differential effects be ascertained with respect to the specific type of instructional treatment (e.g., input practice, output practice)?
3. If differences between treatments can be ascertained, are the differential effects the same across all of the outcome measures? Or are there differences between receptive and productive outcome measures?

Although three research questions were formulated, we assumed—based on existing findings in SLA research—that FFI would indeed have an effect on the participants’ performance with regard to the targeted temporal FMU mappings. Consequently, we did not include a control group to address the first research question directly.

88 See Section 4.2.3.2 for the operationalizations of the treatments.
The research hypotheses (null hypotheses (H₀) and alternative hypotheses (H₁)) that were formulated to accompany the three research questions in Study 1 were the following:

**Research hypotheses**

1. **H₀**: Explicit FFI does not have an effect on Dutch-speaking ESL learners’ performance with regard to the selected complex ESL target features.
   
   **H₁**: Explicit FFI does have an effect on Dutch-speaking ESL learners’ performance with regard to the selected complex ESL target features.

2. **H₀**: No overall differential effects can be ascertained depending on the instructional treatment (input practice, output practice).
   
   **H₁**: Overall differential effects can be ascertained depending on the instructional treatment (input practice, output practice).

3. **H₀**: The differential effects of the explicit FFI are the same across all outcome measures. There are no differences between receptive and productive outcome measures.
   
   **H₁**: The differential effects of the explicit FFI are not the same across all outcome measures. There are differences between receptive and productive outcome measures.

With respect to the hypotheses formulated for the first research question, it was expected that explicit FFI would indeed have an effect on (Dutch-speaking) ESL learners’ test performances as far as the selected complex ESL target features were concerned (see Chapters 1, 2 and 3). For further information on the hypotheses formulated for the first research question, see Chapter 6, Section 6.1.

As far as the second and third research questions were concerned, it was expected that there would be a difference between the varying forms of explicit FFI and the performances on receptive and productive outcome measures. However, since the evidence in the literature on the differential role of receptive and productive practice is contradictory at best (see Chapter 3), a non-directional research hypothesis about the nature of the differential instructional effects was consciously formulated. This hypothesis simply states that differences between input practice and output practice were expected. However, no claims were made about the nature of these differences.
It should be borne in mind that Study 1 was the first of three studies into the effects of explicit FFI. Although Study 1 has not explicitly been referred to as a pilot study, it was regarded as a test run for both the instruments being used for data collection and the various forms of explicit FFI. As a result, methodological and design-related features will also be explicitly discussed in the conclusions and implications sections of this chapter.

4.2 Method
4.2.1 Design
The experimental design for Study 1 was drawn up using an overall four-way 3 (Treatment) x 2 (Time) x 2 (Modality) x 4 (Outcome Measure) mixed-design template. In Study 1, we had one between-subjects variable and three within-subjects variables. The between-subjects variable was treatment, which had three levels (input only, input practice, output practice). The within-subjects variables were time, with two levels (pretest, posttest), modality, with two levels (receptive, productive), and outcome measure, with four levels (grammaticality judgement (GJ), selected response (SR), constrained constructed response (CCR) and translation (TR)). The within-subjects variables were not all hierarchically nested. While modality was hierarchically nested under time, outcome measure was not hierarchically nested under modality since each level of the modality variable had its own two outcome measures. Two of the four outcome measures (GJ and SR) were hierarchically nested under the receptive level of the modality variable, while the two other outcome measures (CCR and TR) were hierarchically nested under the productive level of the modality variable. In principle, it would have been possible to conduct an analysis including all independent variables. However, it was inappropriate to do so in this case, that is, to first conduct a 3 x 2 x 2 x 4 analysis because there were not four outcome measures in each modality; two of the outcome measures were found in the receptive modality and the two other outcome measures were found in the productive modality. Therefore, we consciously decided to carry out two separate analyses (see Section 4.3).
dependent variable

1
continuous

written performance on a 24-item tense-related test consisting of four separate 6-item outcome measures of either a receptive or a productive nature: (1) grammaticality judgement, (2) selected response, (3) constrained constructed response and (4) translation

independent variables

4
categorical

1 between groups

3 within groups

Between-groups variable:
1) Treatment (3 levels)
   1) Input only
   2) Input practice
   3) Output practice

Within-groups variables:
1) Time (2 levels)
   1) Pretest
   2) Posttest
2) Modality (2 levels)
   1) Receptive
   2) Productive
3) Outcome measure (4 levels)
   1) Grammaticality judgement
   2) Selected response
   3) Constrained constructed response
   4) Translation

Figure 4.1. Design box for Study 1 (3 x 2 x 2 x 4 mixed design)

4.2.2 Participants
The participating ESL learners were students who had all enrolled in a three-year translation programme in Antwerp, the second largest city in Belgium located in the north of the country, where the official language is Dutch. The three-year translation programme would eventually lead to a bachelor’s degree in applied linguistics with a specialization in translation.

Attrition. Although initially 88 participants took part in the pretest, the number of participants dropped to 77 during the six experimental sessions. This 12.5% drop was the result of various extraneous factors (e.g., illness, participants’ forgetfulness, schedule changes). For the discussion of the experimental data, the attrition rate was taken into consideration. In addition, seven participants had a total pretest score of 21 or higher on a possible maximum
score of 24. In other words, these seven participants answered more than 85% of the pretest items correctly. Since these participants performed near-ceiling in the pretest, we took them to be familiar with the target structures under investigation. Consequently, instruction was unlikely to improve their performance in any significant way. Only the characteristics of and the data from the remaining 70 participants were considered in the descriptive and inferential data analyses, which were conducted post-experimentally. Any information provided below takes the 70 participants as the 100% benchmark.

Assignment to treatment groups. The participants were requested to register via e-mail before the first experimental session. After the registration deadline, the participants were randomly assigned to one of the three treatment groups as follows: 24 (input only), 21 (input practice), 25 (output practice).

ESL grammar course. The experimental sessions were conducted as part of the participants’ regular English Grammar course in the first year of their English bachelor programme. The participants were briefed about the course expectations and the general setup of the English Grammar course during the first English Grammar class at the beginning of the academic year 2007-2008. English grammar instruction amounted to two hours of instruction on a weekly basis and the course was taught by the author of this doctoral dissertation throughout the entire academic year. One hour of the course was dedicated to theoretical instruction, which was taught to one large group consisting of all the participants enrolled for the English Grammar course. The remaining hour was dedicated to practical exercises, which were taught to smaller groups of approximately 25 to 30 participants. Because of both the experimental nature of the sessions and the apparatus and materials required during the sessions, all of the sessions took place during the participants’ exercises class.

ESL language proficiency. The participants were required to have pre-existing knowledge of English. For the English Grammar course, the participants’ level of proficiency was expected to reach at least level B1 of the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001). On the whole, the participants’ overall level of proficiency was categorized as intermediate to upper intermediate though no prior testing was carried out. The categorization was based on a combination of the English Grammar course requirements and the Belgian secondary school goals (also known as eindtermen in Dutch) for English for pupils in their last year of general secondary education, which are set at B1 of the CEFR.

Academic exposure to ESL English. The participants received either six or seven hours of English language classes on a weekly basis. All of the courses were taught throughout the entire academic year. The total amount of ESL instruction depended on whether the participants had chosen English as a first/second foreign language or as a third foreign language in their foreign language combinations. For the participants with six hours of English, the English courses in the bachelor programme consisted of the following three ESL
STUDY 1

courses: (1) English Grammar (two hours per week), (2) English Practicals (two hours per week) and (3) General Translation: English into Dutch (two hours per week). The participants with seven hours of English had the same three courses referred to above and an additional English Texts course (one hour per week).

Academic exposure to other L2s. The foreign language combinations were completely dependent on the participants’ own choices within the constraints of the bachelor programme structure for which they had enrolled. Depending on which year of the bachelor programme that the participants were in, they were expected to select two or three L2s (in addition to Dutch, which was considered their L1). Not all of the participants were first-year students but all of the participants were taking English courses in the first year. Regular first-year participants were expected to select two L2s. Second-year participants were able to select English as a third L2, which was then added to their existing combination of two L2s. In other words, although second-year participants were taking foreign language courses in the second year of their bachelor programmes, they were all taking English language courses in the first year. The languages in the bachelor programme were divided into three groups: (1) A language, (2) B languages and (3) C languages (see Table 4.1). The participants were required to select one of the following language combinations: A-B-B or A-B-C (for regular first-year participants) and A-B-B-B, A-B-C-B, A-B-C-C (for second-year participants with a third optional L2).

| Table 4.1. Language groups (bachelor in applied linguistics with a specialization in translation) |
|----------------------------------------|----------------------------------------|----------------------------------------|
| A language | B languages | C languages                               |
| Dutch      | English     | Chinese (Mandarin)                        |
|            | French      | Italian                                   |
|            | German      | Portuguese                                |
|            |             | Russian                                   |
|            |             | Spanish                                   |

Of the 70 participants, 58 had chosen English as part of an L2 combination consisting of two L2s. This meant that these 58 participants received seven hours of ESL instruction on a weekly basis. In addition, they had one other L2 in their L2 combination, for which they also received seven hours of instruction on a weekly basis. Furthermore, these participants were also expected to take Dutch language courses, since Dutch was their A language. The remaining 12 participants had chosen English as part of an L2 combination consisting of three L2s. These participants received either six or seven hours of English on a weekly basis, which they combined with two other L2s and Dutch (as their A
language). Table 4.2 provides a detailed overview of the participants’ various L2 combinations.

<table>
<thead>
<tr>
<th>English</th>
<th>other L2(s)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>Spanish</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Chinese (Mandarin)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Russian</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Italian</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>French</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Portuguese</td>
<td>1</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>58</td>
</tr>
</tbody>
</table>

| yes     | French and Italian | 3  |
|         | French and Spanish | 2  |
|         | Chinese and Portuguese | 1 |
|         | French and German  | 1  |
|         | French and Portuguese | 1 |
|         | French and Russian | 1  |
|         | German and Spanish | 1  |
|         | Italian and Spanish | 1 |
|         | Portuguese and Spanish | 1 |
| Subtotal|                | 12 |
| Overall total |            | 70 |

**Gender and age.** Of the 70 participants, 10 were male and 60 were female. In addition, the participants ranged in age from 18 to 45 years and the mean age was 21 years ($SD = 4$, mode = 19, median = 20).

**L1 backgrounds.** The participants were asked about their L1 backgrounds and more specifically about which language(s) that they themselves regarded as their L1(s). Of the 70 participants, 69 indicated that they had one L1. Of those 69 participants, 64 participants had Dutch as their L1. In alphabetical order, the remaining five participants’ L1s were distributed as follows (with the number of participants in parentheses): German (1), Polish (1) and Russian (3). Only 1 participant indicated that she had two L1s, which were Dutch and English. It should be pointed out that, because of both the specific language-oriented nature of the bachelor programme and the focus on translation into Dutch, incoming bachelor students with degrees of secondary education issued by schools in countries where Dutch was not the official language had passed a Dutch proficiency test (level C1 of the CEFR) before being allowed to enrol for the bachelor programme.
Remuneration. Since the experimental sessions were carried out during regular English Grammar classes, the participants did not receive any form of financial remuneration for their participation in the sessions. They did, however, receive a final mark, which was a combined score of the pretest and posttest results. This mark was made known to the participants a couple of weeks after the experimental sessions had been completed. The participants were informed in advance that participation in the study could be beneficial to them in the form of possible bonus credit for the English Grammar course. They were also explicitly told that any marks awarded could in no way contribute negatively to their final mark for the English Grammar course. The bonus credit was simply a means to motivate the participants to take part in the experimental sessions.

4.2.3 Apparatus and materials
4.2.3.1 Apparatus
The participants completed all of the six experimental sessions in the same computer language lab, in which each participant had a desktop computer to work on. All of the experimental sessions were programmed by a computer specialist, who had also received formal language training in higher education. The sessions were subsequently hosted on an external SQL server. The decision to host the sessions externally was a conscious one since external hosting not only allowed for possible work on, manipulation of and updates to the experimental sessions to be carried out remotely but also created a technically stable and reliable environment in which the experimental sessions could be carried out and monitored non-invasively. Technical support was available throughout the entire duration of the six experimental sessions but no major computer-related or network-related problems which could possibly have jeopardized the execution of the experimental sessions were reported. Minor technical problems were occasionally reported but these problems were of a temporary nature and did not jeopardize the experimental sessions. An example of a minor technical problem was the participants’ forgetfulness with respect to their network usernames and passwords. This problem was anticipated and circumvented by creating backup accounts in advance, with which affected participants were able to log on.

4.2.3.2 Materials
Apart from the materials made available to them on-screen, the participants did not use any other materials such as course notes, paper and writing material. The ESL target structures under investigation in the various experimental sessions were the past and the present perfect when used to locate bygone situations in standard, formal present-day English (see Chapters 1, 2 and 3). Taking into account the experimental design and the instructional approach, I have subdivided this section into the following subsections: pretest, treatment the-
ory, treatment practice and posttest. This subdivision reflects the setup of the experimental sessions in Study 1.

Pretest. In light of the ESL target structures under investigation, a computerized 24-item pretest was administered over two weeks (12 items per week). The majority of the texts that were presented to the participants were taken from the Dutch popular-scientific magazine *Quest*, which appears on a monthly basis. In addition to the regular multithematic (monthly) publications, the magazine is also occasionally published in thematic format (with popular topics such as history, science, travel) at larger time intervals. The thematic publications appear in addition to the regular publications and for the experiment both types of publication were used. The texts were selected on the basis of possible problems with respect to the instantiations of the complex ESL target features under investigation, that is, the complex past/present perfect distinction when locating bygone situations in present-day English. The initial reason for choosing Dutch texts was that the participants would be asked to translate complete texts for one of their outcome-measure. However, extraneous variables such as processing requirements and time restraints led to the selection of sentences in the texts, which the participants then had to translate. Since a corpus with a variety of texts had already been drawn up when this decision was taken, the translated texts were used since they were perfectly suited as functionally equivalent English texts of the original Dutch source texts. The Dutch source texts were translated by a native speaker of English with a master’s degree in applied linguistics and a specialization in translation (Dutch-English-German). In addition, the translator had at least ten years of experience in translating from Dutch and German into English. The translated texts were proofread and edited—wherever necessary—to ensure that they were all consistent with respect to presentation format, (overall) text difficulty and text length. Subsequently, the texts with the accompanying questions were distributed among a panel of translators and proofreaders to ensure that acceptable levels of (content) validity were safeguarded.

The pretest and posttest were designed to measure the Dutch-speaking ESL learners’ ability to comprehend and produce the past and the present perfect when used to locate bygone situations in present-day English. Since the experimental focus was on both the comprehension and the production of these two tenses, it was deemed necessary that both comprehension-related and production-related features were made available in the tests. Consequently, four tense-related outcome measures were selected, which reflected the comprehension/production dichotomy (cf. below).  

---

89 See www.quest.nl for additional information about the format and contents of the magazine.
In addition to selecting the four outcome measures, it was important to select texts for which a professional (linguistic) consensus was available with regard to the use of the past and the present perfect. To achieve this consensus a corpus of 60 texts was initially compiled. Subsequently, the corpus texts were grouped into 4 categories taking into account the four outcome measures. This resulted in 15 texts for every outcome measure. Using the four task types discussed below, tense-related questions were drawn up for the 60 texts. Subsequently, the corpus was sent to a panel of seven experienced translators. The selected translators were either native speaker or near-native speakers of English. All of the translators had at least five years of experience with translating both general texts and more specialized texts (e.g., contracts, press releases, manuals) from Dutch into English. The panel members were asked to answer the tense-related questions to the best of their abilities and to provide feedback about both general text features (e.g., lexical text complexity, structural text complexity, topic complexity) and specific ESL target feature characteristics (e.g., FMU mapping features, frequency, irregularity) in the context of the pretest and posttest sessions. The answers and the feedback provided by the panel members were compared. Based on the feedback, a total of 44 texts were selected from the corpus of 60 texts. The choice of these 44 texts was based solely on the level of consensus found in the panel members’ replies and feedback. The remaining texts (16 texts) were investigated and the answers and feedback received from the panel members for those texts were taken into consideration when optimizing the texts and making them acceptable and consistent with the 44 texts already accepted. Some of the remaining texts were discarded because in the end they were not deemed useful for the research purposes. Every effort was made to ensure that the texts were consistent with respect to both the ESL target features under investigation and the (linguistic) contexts in which the ESL target features were found.

In addition to receiving a tense-related question for every text during the pretest, the participants also received three additional questions related to other aspects of the texts (e.g., cultural references, grammar items not related to tense, lexis). This decision was taken to ensure that the participants would interact more thoroughly with the contents of the texts and to try to get the participants to focus on the text as a structured sample of extended discourse.

As stated above, the four tense-related outcome measures that were used in the pretest consisted of two receptive outcome measures (grammaticality judgement, selected response) and two productive outcome measures (constrained constructed response, translation). The numbers of test items for every outcome measure were calculated by simply dividing the total number of pretest items by the number of outcome measures. This resulted in six items for every outcome measure in Study 1.

The first type of receptive outcome measure was a form of grammaticality judgement (GJ) task type. For the questions related to this type of outcome
measure the participants were asked to evaluate the grammaticality of a finite verb form—either a past or a present-perfect verb form—which appeared in bold in a (linguistic) context after the text had appeared on-screen for 40 seconds. One finite verb form per text appeared in bold and for that specific verb form the participants were asked to select whether the verb form was grammatically correct or grammatically incorrect in standard, formal present-day English. The participants were able to see the verb form in its (linguistic) context whilst thinking about their answers. The participants made their choices by selecting the option which they believed to be correct. Participants were free to use British English or American English as their points of reference. Because of this choice that they had, all the questions related to all of the outcome measures were asked in as neutral a way as possible with regard to regional varieties of English. An example of a grammaticality judgement text and task type may be found in Figure 4.2.

Bees are social animals. If one of them finds a food source, it tells its kind. But how? By dancing! Wagging the back part of its body, the happy finder runs an eight. And it is the stripe which is most in the middle of the eight which indicates the right direction. Bees also pay attention to the dancer’s wagging speed since it says something about the distance of the food source. It is with these wonderful observations that the Austrian biologist Karl von Frisch has won the Nobel Prize in 1973. However, many of his colleagues remained sceptical.

Question
In standard, formal present-day English, the verb form in bold in the text above is considered … .

- grammatically correct
- grammatically incorrect

Figure 4.2. Grammaticality judgment (GJ): Example text and task type

The second type of receptive outcome measure was a form of selected response (SR) task type. The participants were asked to select the finite verb form which—according to them—was the grammatically correct verb form for the non-finite verb which appeared in bold in the (linguistic) context provided. The possible answers contained the key (either a past verb form or a present-perfect verb form depending on, for example, the linguistic context) and two distracters (either a past verb form or a present-perfect verb form to complement the key and an additional ungrammatical finite verb form). An example of a selected response text and task type may be found in Figure 4.3.
It is still some sort of Holy Grail in artificial intelligence research: the Turing test. Alan Turing (1912-1954) came up with this test in 1950 to determine whether a computer is truly intelligent. Have a test subject communicate with a computer and a human being, using a keyboard and a monitor. If, after a couple of minutes, the researcher is unable to determine who the human being is or who the computer is, the computer can be considered intelligent. To date, no one has been able to develop a computer which passes this test convincingly.

Question
Select the appropriate verb form for the verb in bold in the text above to make the sentence acceptable in standard, formal present-day English.

• has come up     • came up     • will have come up

The first type of productive outcome measure was a form of constrained constructed response (CCR) task type. The participants were asked to provide a finite verb form which—according to them—was a grammatical verb form for the non-finite verb which appeared in bold in the (linguistic) context provided. An example of a constrained constructed response text and task type may be found in Figure 4.4.
CHAPTER 4

From a smelly barrel behind a door with a heart on it to the shiny white model we know today: the toilet ... *(to undergo)* a huge series of developments. But the white porcelain object could be regarded as somewhat boring. This is what Meike van Schijndel must have been thinking when she designed the urinal in the shape of a woman’s mouth. Sadly, not everyone was as enthusiastic as she was about her revolutionary idea for the urinal. When her idea was presented in the United States in 2004, it was met with protests.

Question
Put the verb in bold in an appropriate form to make it acceptable in standard, formal present-day English.

Figure 4.4. Constrained constructed response (CCR): Example text and task type

The second type of productive outcome measure was a form of translation (TR) task type. This outcome measure was selected because of the participants’ specific academic environment. Although most of the participants were only in the first year of their bachelor programme in applied linguistics (with a specialization in translation), the decision was taken to have the participants translate one sentence per text as opposed to translating entire texts. Even though the selected texts were structurally and translationally relatively straightforward from a professional translator’s point of view, the participants would probably have still had to use specific translation strategies if they had been asked to translate complete texts in the allotted time frames. In turn, the required translation strategies may have caused additional problems (e.g., increased processing costs, time pressure), which the participants would have experienced as too demanding with their limited experience with translating complete texts into English. In an attempt to avoid the effects of such extraneous variables, the one-sentence translation decision appeared to strike a happy medium. For the questions related to this type of outcome measure, the participants were asked to translate into standard, formal present-day English a Dutch sentence in bold in the (linguistic) context provided. An example of a translation text and task type may be found in Figure 4.5.
Eeuwenlang hebben wetenschappers gedacht dat de aarde uniek was. But Nicolaus Copernicus showed us that the Earth was simply one of several planets orbiting the sun. Back then, astronomers thought that the sun was unique, but it became clear fairly quickly that the sun was one of countless stars in the Milky Way galaxy. After that, it was generally assumed that the Milky Way galaxy was unique but that galaxy turned out to be one of billions of galaxies in the universe. So why should we now assume that our universe is unique? In the past, science fiction authors have written a lot about parallel universes but during the last years such universes have also been discussed in scientific journals.

Question
Translate the sentence in bold in the text above into standard, formal present-day English.

Treatment theory. After completing the pretest, the participants received an on-screen theory session, during which all of the participants were given the exact same theoretical information about both the problem of choosing between the past and the present perfect to locate bygone situations and the most important factors which influence the choice between the past and the present perfect in such situations. Example slides may be found in the appendix.

The first slides contained general information and focused on explaining metalinguistic concepts such as situation, (non-)bygone situation, linguistic context, pragmatic context. After defining the relevant concepts, the participants were presented with some observations about the use of the past and the present perfect when they are used to locate bygone situations in Dutch and in English. Some of these observations were consciously drawn up to function as observations about contrastive uses of these two tenses (Dutch versus English) in an attempt to address possible processing problems that many of the participants experienced. Subsequently, the participants received additional slides containing more specific contrastive information (with example sentences in Dutch and English) on the importance of context, the creation of past-time contexts (linguistic and pragmatic) and the use of past-time indicators (e.g., time adverbials, place adverbials) in creating such contexts.
After presenting the participants with these elements at a sentential level, the focus was placed on the suprasentential level (i.e., extended discourse). In this respect, the participants were presented with short, coherent texts—similar in text difficulty and text length to the test items—which were aimed at unifying the theoretical information already presented and at providing the participants with concrete instances of theory-related contrastive texts (Dutch versus English). These texts were followed by a summary of the items discussed and by a few examples of sentences in which either no explicit context (linguistic or pragmatic) was provided or the (linguistic) context provided was one created with multi-zone adverbials (thus allowing both the past and the present perfect but with various readings). The final slide of the theory session was a flow-chart summary of the most relevant theoretical information presented during the theory session. In essence, the summary was an on-screen multi-step plan for deciding which tense (past or present perfect) should be used when locating bygone situations. This plan was a summary of the main theory items from the theory session. The theory session was identical for all three treatments (input only, input practice, output practice) and lasted one instructional session (55-60 minutes). However, since the participants were free to navigate through the theory slides at their own pace, not all of the participants needed the full instructional session.

**Treatment practice.** There were three experimental groups (input only, input practice, output practice), which all received different forms of tense-related treatment practice. The materials that were used for the practice sessions consisted of seven texts, which were identical for all three instructional treatments and which were similar to the texts that the participants had received in the pretest. In other words, all three treatments received the exact same seven texts during the treatment practice sessions. The participants in the input-only group, which received a relatively simple yet straightforward form of input enhancement, were presented a complete text. After the text had appeared on-screen for 40 seconds the verb form of interest for the treatment session (i.e., either a past or a present perfect) changed colour (from black to blue) and appeared in bold in the text. In addition, other relevant temporal information present in the text (e.g., temporal adverbials) also changed colour (from black to red). The input enhancement was completed by means of a ‘callout’ shape in the form of a balloon, which brought together the verb form and the other relevant temporal information highlighted in the text. As such, the participants in the input-only group were given simple instructions to click their way through the slides. No further instructions were provided to the participants in the input-only group during these sessions. An example of such input enhancement may be found in Figure 4.6.
Louis Armstrong (1901-1971) was one of the greatest jazz trumpet players of the 20th century, but he ended his career though as a singer. Lips are not suited to vibrate against the mouthpiece of a trumpet day in and day out. Armstrong struggled with this fact all his life. Even at the height of his career in the 1930s, his lips had been ruined as a result of his energetic blowing technique. Every time he put his instrument against his lips it felt as though a scorching poker was being held against them. At the end of a concert he was forced to swallow the constant stream of blood. There was a concert in 1935 during which the blood flowed down his shirt during his performance. In the end, Armstrong was forced to cancel so many performances that his manager sued him for breach of contract.

The input-practice group received the same texts as the input-only group but with a different form of treatment, which was comprehension-based. The first part of the treatment was identical with the treatment that the input-only group received. However, in addition to the input enhancement visible in Figure 4.6, the input-practice group received an additional input-related question. The participants were asked to select one of two pictures which they believed represented the temporal relations expressed by the highlighted finite verb form and the other relevant temporal information selected for input enhancement (e.g., temporal adverbials). Upon confirmation of their choices, the input-practice participants received a slide containing metalinguistic feedback on the linguistic context in which both the finite verb form and the other relevant temporal information selected for input enhancement were found. The feedback provided the input-practice participants with the correct verb form and was provided regardless of whether the participants’ answers were correct or not (see Figure 4.7). In addition, the input-practice participants also received the picture which correctly represented the temporal relations.
CHAPTER 4

Question
Which of the following pictures reflects the time frame referred to by the verb in bold/blue in the text above?

Answer

Feedback
A linguistic context can be found in the text and it is the linguistic context which contains a past-time indicator, which, in turn, determines the use of the past tense. The past-time indicator is the reference 1901-1971, which can be found in a sentence preceding the tense form and the indicator is actually referred to in the sentence in which the verb form appears using the expression all his life. This means that the relevant linguistic context is not just in the rest of the sentence (i.e., the sentence minus the tense form) but also in a sentence preceding the tense form. The adverbial 1901-1971 makes it clear that (a) we are talking about a bygone situation and (b) there is no focus on ‘now’ as far as Armstrong’s life and his ‘struggle’ are concerned. The focus is on ‘then’ and so no link is available with ‘now’.

Figure 4.7. Picture selection task with feedback for input-practice group

The output-practice group received the same texts as the input-only and input-practice groups but with productive practice in the form of translation taking into account the participants’ academic environment. This resulted in texts in
which most of the input was in English. However, the sentences which the output-practice participants were asked to translate were in Dutch. Using the example text provided above this led to the form of production-based practice for the output-practice participants represented in Figure 4.8.

***

Louis Armstrong (1901-1971) was one of the greatest jazz trumpet players of the 20th century, but he ended his career though as a singer. Lips are not suited to vibrate against the mouthpiece of a trumpet day in and day out.

Armstrong heeft zijn leven lang met dit gegeven geworsteld. Even at the height of his career in the 1930s, his lips had been ruined as a result of his energetic blowing technique. He was always treating his injuries and blisters using cream. Every time he put his instrument against his lips it felt as though a scorching poker was being held against them. At the end of a concert he was forced to swallow the constant stream of blood. There was a concert in 1935 during which the blood flowed down his shirt during his performance. In the end, Armstrong was forced to cancel so many performances that his manager sued him for breach of contract.

Translate the sentence in bold in the text above into standard, formal present-day English.

Feedback

A linguistic context can be found in the text and it is the linguistic context which contains a past-time indicator which, in turn, determines the use of the past tense. The past-time indicator is the reference 1901-1971, which can be found in a sentence preceding the tense form and the indicator is actually referred to in the sentence in which the verb form appears using the expression all his life. This means that the relevant linguistic context is not just in the rest of the sentence (i.e., the sentence minus the tense form) but also in a sentence preceding the tense form. The adverbial 1901-1971 makes it clear that (a) we are talking about a bygone situation and (b) there is no focus on 'now' as far as Armstrong's life and his 'struggle' are concerned. The focus is on 'then' and so no link is available with 'now'.

A possible English translation is then: Armstrong struggled with this all his life and not *Armstrong has struggled with this all his life.

Figure 4.8. Example text and translation task for output-practice group
Upon confirmation of their translations, the output-practice participants received feedback on their translations. This feedback was identical with the metalinguistic feedback provided to the input-practice participants (see Figure 4.8).

**Posttest.** A week after completing the treatment sessions, the participants were asked to take part in an unannounced immediate posttest. The setup of the posttest was identical with the setup of the pretest. The only difference was the actual nature of the 24 posttest items. Whereas the pretest consisted of 24 new items, which the participants had not seen before, the posttest consisted of twelve ‘old’ items (i.e., items which had been recycled from the four pretest outcome measures, with three recycled items per outcome measure) and twelve ‘new’ items, which were new for all of the participants. The twelve ‘old’ items were not randomly selected but were chosen based largely on item difficulty indices calculated after analysing the pretest answers.

### 4.2.4 Procedure

Study 1 took place in the months of April and May of the academic year 2007-2008 and consisted of six experimental sessions, which were divided as follows: pretest (two sessions), treatment theory (one session), treatment practice (one session), posttest (two sessions).

The participants were requested to enrol before the beginning of the first experimental session so that user accounts could be created in advance for all of the participants. The participants were told that the experimental sessions would focus on an aspect of English grammar but no further information was provided to prevent the participants from preparing for the experimental sessions. In addition, the participants were not aware in advance of the contents of the individual experimental sessions.

During the first experimental session, the participants were informed that they would be taking a test and that all the required information would be provided on-screen. After all of the participants had logged on to their accounts, they started the pretest and were given on-screen instructions before starting the actual test. Once the participants had finished reading the pretest guidelines and instructions, the key features of the guidelines and instructions were summarized in class by the instructor. Any remaining questions about the instructions that the participants had were answered in all of the groups in as consistently a way as possible. Backtracking features were not made available to the participants during the pretest sessions. The participants were warned about

---

91 Strictly speaking, the posttest was not immediate since there was one or two weeks between the last instructional session and the first or second posttest session. However, since the experimental sessions were class-bound, I will assume that one week was the standard period between sessions and that the posttest followed the normal course of the experimental sessions. In Study 3, for example, a second posttest was administered eleven weeks after the first posttest, with the first posttest regarded as the immediate posttest and the second posttest as the delayed posttest.
the absence of such features before the pretest started. By making them aware of this in advance, they could ensure that their pretest answers were their final answers before continuing with the next pretest item. The 24 pretest items were presented to the participants in context, namely in the form of relatively short, coherent texts (see the outcome measures above for examples of the texts and task types). The order of the texts was randomized and this randomization factor was programmed into the relevant software script, which resulted in unique orders for every participant. Each text came with four questions (one tense-related question and three questions about other text features such as cultural items, lexical items). After completing twelve texts the participants were given on-screen instructions to log off. As far as the participants were aware this was the end of the pretest. However, one week later the participants were given on-screen instructions to complete the pretest and at the end of the second experimental session all the participants had completed the entire pretest, which consisted of 24 items. The pretest had to be split because of time constraints.

One week later, the participants received explicit information about the ESL target features under investigation: the past and the present perfect when used to locate bygone situations in present-day English. The participants had one hour to complete the theory session. In practice, many of the participants did not require an entire hour to complete the treatment theory session but to avoid any effects of extraneous factors such as time stress, it was decided that one hour should be made available to all of the participants for this session. Instructions were once again provided on-screen before the treatment theory session commenced. Once the participants had finished reading the guidelines and instructions, the key features of the guidelines and instructions were summarized in class by the instructor and any remaining questions that the participants had were once again answered. Subsequently, the participants were free to navigate through the treatment theory session by pressing the required buttons on-screen. This time, backtracking features were made available for the theory session and the participants were made aware of the presence of these features. The theory session consisted of 20 MS PowerPoint slides (see description above), which were imported into the programming software which was used for the experimental sessions.

One week after receiving the theory session, the participants took part in the practice session. As was the case with the other sessions the participants received instructions and guidelines on-screen. Backtracking features were made available and the participants were made aware of these features before the session started.

The final two experimental sessions consisted of a posttest, which was set up in exactly the same way as the pretest. The only difference was the nature of the 24 items on the posttest. Of the 24 items, twelve had been recycled from the pretest based on the item difficulty indices, which had been calculated after
the pretest data had been coded. Once again, the participants required two experimental sessions to complete the entire posttest. At the end of the last experimental sessions, the participants were told that they had completed all of the planned experimental sessions.

4.2.5 Coding and scoring

Both the pretest and the posttest were completely handscored. All 24 items on each of the tests were scored as either right or wrong with right answers being awarded a ‘1’ score and wrong answers a ‘0’ score. The coding of answers focused on meaning-related and use-related features of the temporal ESL target features. However, indirectly, form-related features of the temporal FMU mappings had to be used as indicators of the targeted finite verb forms. Since complete formal accuracy of the FMU features was not one of the goals of the treatments, the scoring of the formal features was performed more leniently. This resulted in ungrammatical past forms such as *he promised and ungrammatical present perfect forms such as *the problem has arised, *he have came being marked as correct as long as the participants showed that they knew which tense to use even though they produced formally ungrammatical verb forms as a result of orthographical and/or grammatical problems (e.g., number agreement). The problem of formal inaccuracy was an issue especially for the productive outcome measures, for which participants were asked to produce either individual grammatical verb forms or grammatical verb forms as part of translations of entire sentences. As far as the translation outcome measure was concerned, a decision was taken not to score the entire translations provided by the participants. Instead, the targeted finite verb forms were analysed with respect to the meaning-related and use-related features of the past/present perfect distinction.

In theory, a system of partial-credit marking could have been used (especially for the productive outcome measures). However, such marking would have complicated the analyses of the data in several ways. First, it would have awarded proportionally higher scores to specific test items without any way of feasibly teasing apart the various features of the FMU mappings in the answers. Second, partial-credit marking would have led—especially for the translation outcome measure—to the scoring of data (e.g., lexis, syntax) irrelevant for the research being carried out. In turn, this would have led to an imbalance in the coding of the various outcome measures, which would have provided a distorted picture of the results. The focus was purely on the targeted finite verb forms and to keep the marking consistent across all of the four outcome measures, the same coding and scoring procedures were used for all of the outcome measures.
4.2.6 Statistical analyses

As already mentioned in Section 4.2.1, it would have been possible to conduct an analysis including all independent variables. However, considering the dependency of the outcome measure variable on the modality variable, it was inappropriate to do so in this case, that is, to conduct a 3 (Treatment) x 2 (Time) x 2 (Modality) x 4 (Outcome Measure) analysis. Consequently, we decided to run two separate analyses (see Section 4.3).

The assumptions which underlie repeated-measures analyses of variance (RM ANOVAs) were checked. Negative skewness (and, by extension, non-normal distribution) was a common feature of the data. Transformation of the data was occasionally required to meet the assumption of normality of data, which underlies the RM ANOVAs used for interpreting the data. Because of the (moderate) negative skewness, the square root transformation \( \sqrt{Z-X} \) was occasionally used following suggestions from Tabachnik and Fidell (2007). Analyses were carried out with transformed and untransformed data but the results were the same. The raw data were checked first and any violations were investigated with and without transformations. Although the analyses were performed at times on transformed data, all figures and tables presented below contain the raw data from Study 1. Because of the nature of the transformation (taking the square root of the reversed data), inclusion of the transformed data in the tables would have unnecessarily complicated any interpretations.

In addition to the assumption of normality of data, other assumptions of RM ANOVAs (e.g., homogeneity of variances, sphericity) were also explored at the three levels of statistical analysis. Of special importance in this respect was the assumption of sphericity. Field (2009) defines sphericity as “a more general condition of compound symmetry” (p. 459). Larson-Hall (2010) defines the concept of sphericity as the concept which “measures whether differences between the variances of a single participant’s data are equal” (p. 336). Since sphericity is an issue only when the repeated-measures variable has more than two levels, no references to this RM ANOVA assumption will be made for Study 1 since the repeated-measures variable time consisted of only two levels (pretest, posttest).

Outliers were defined for each cell in the experimental design as test scores which were located 1.5 times the interquartile range below or above the first and third quartiles respectively. The outliers that were discovered were investigated and analyses were run with and without values defined as outliers. Deal-

---

In the square root transformation \( \sqrt{Z-X} \), the X value represents the value(s) of the variable to be transformed and the Z value represents a constant which is used to ensure that \((Z-X)\) does not equal 0. For the analyses in experiment 1, the possible constants were 25 (24+1), 13 (12 + 1) and 7 (6+1) for the three levels of statistical analysis (see Figure 4.9). For Study 1, the choice between the three constants depended on whether the variable was the sum of all four outcome measure for the complete tests (24 items), the sum of two outcome measures (twelve receptive items, twelve productive items) or one individual outcome measure (six items).
ing with outliers is often problematic (Field, 2009; Larson-Hall, 2010) so a decision was taken to include the outliers in the analyses since they did not significantly affect the statistical analyses.

A significance level of .05 was set for all statistical tests unless otherwise indicated. In addition, effect size was assessed using the partial eta squared ($\eta^2_p$) measure, with .01, .06 and .14 as standards for small, medium and large effect sizes respectively (Huck, 2009).

### 4.2.7 Validity and reliability

The measures that were taken to safeguard (content) validity were described in Section 4.2.3.2. In addition to investigating validity, reliability was also an issue which was taken into account. Reliability (in the form of internal consistency) was measured using Cronbach’s alpha coefficient ($\alpha$). Although the views about acceptable Cronbach’s alpha values are not always in agreement, it is generally accepted that a value of .70 or higher is acceptable. The Cronbach’s alpha coefficients for the pretest and posttest in Study 1 were .46 and .76 respectively for all of the items as a whole. Although the pretest shows relatively low reliability, it would be unwise to dismiss the reliability of the pretest altogether based on this low Cronbach’s alpha coefficient. Huck (2009) states the problem of low validity and reliability coefficients as follows: “[K]eep in mind that validity (like reliability) is really a characteristic of the data produced by a measuring instrument and not a characteristic of the measuring instrument itself” (p. 95). This, of course, is true of the reliability of the pretest. If one bears in mind that we selected a complex ESL target feature for instruction, with which ESL learners of all levels of proficiency experience problems, it is not a complete surprise that low(er) Cronbach’s alpha coefficients were found. What is more telling is the increase in the Cronbach’s alpha coefficients from pretest to posttest, which could partly be explained by an increase in receptive and productive awareness of the complex past/present perfect distinction.

### 4.3 Quantitative results

The results reported in this section are of a quantitative nature. However, some qualitative data were also collected, which will be reported in a separate section (Section 4.6).

---

93 The standards for small, medium and large effect sizes (using $\eta^2_p$) are general standards and may need to be changed depending on the research design that is being used and the research goals that have been set.

94 See Larson-Hall (2010) for a more detailed discussion of Cronbach’s alpha coefficients and the problems related to evaluating such coefficients.

95 The increase in Cronbach’s alpha coefficients may be partly explained by an increase in the command of the complex past/present perfect distinction but other factors may also have played a role here (e.g., the nature of the pretest/posttest items, the treatments).
Analyses for both main effects and interaction effects were carried out at the three levels of the experimental design as indicated in Figure 4.9.

<table>
<thead>
<tr>
<th>level</th>
<th>items</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>all items</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>receptive items</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>productive items</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>grammaticality judgement items</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>selected response items</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>constrained constructed response items</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>translation items</td>
</tr>
</tbody>
</table>

Figure 4.9. Levels of statistical analysis

As mentioned before, seven students had a total pretest score of 21 or higher, that is, they answered more than 85 per cent of the pretest questions correctly. Since these students performed near ceiling in the pretest, we took them to be familiar with the target structures under investigation. In other words, instruction was unlikely to improve their performance significantly. We therefore conducted the analyses on only those participants with pretest scores of 20 or lower, leaving 24, 21 and 25 participants in the input-only, input-practice, and output-practice groups respectively. We first conducted a three-way 3 (Treatment) x 2 (Time) x 2 (Modality) RM ANOVA, with treatment as the between-subjects variable and time and modality as the within-subjects variables. The descriptives are shown numerically in Table 4.3 and graphically in Figures 4.10, 4.11 and 4.12.
### Table 4.3. Performance on receptive and productive pretest and posttest items by treatment group (maximum scores of 12 on receptive and productive items and 24 on the totality of pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Modality</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI_{95%}</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Receptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input only</td>
<td>24</td>
<td>8.92</td>
<td>1.89</td>
<td>8.12</td>
<td>9.71</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>8.00</td>
<td>1.64</td>
<td>7.25</td>
<td>8.75</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>8.20</td>
<td>1.68</td>
<td>7.51</td>
<td>8.89</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>70</td>
<td>8.39</td>
<td>1.76</td>
<td>7.97</td>
<td>8.81</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input only</td>
<td>24</td>
<td>7.75</td>
<td>1.65</td>
<td>7.05</td>
<td>8.45</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>7.52</td>
<td>2.09</td>
<td>6.57</td>
<td>8.47</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>8.44</td>
<td>1.39</td>
<td>7.87</td>
<td>9.01</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>70</td>
<td>7.93</td>
<td>1.73</td>
<td>7.52</td>
<td>8.34</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Receptive &amp; productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input only</td>
<td>24</td>
<td>16.67</td>
<td>2.94</td>
<td>15.42</td>
<td>17.91</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>15.52</td>
<td>3.11</td>
<td>14.11</td>
<td>16.94</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>16.64</td>
<td>2.53</td>
<td>15.60</td>
<td>17.68</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>70</td>
<td>16.31</td>
<td>2.86</td>
<td>15.63</td>
<td>17.00</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input only</td>
<td>24</td>
<td>8.83</td>
<td>2.37</td>
<td>7.83</td>
<td>9.83</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>8.86</td>
<td>2.82</td>
<td>7.58</td>
<td>10.14</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>8.64</td>
<td>2.56</td>
<td>7.58</td>
<td>9.70</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>70</td>
<td>8.77</td>
<td>2.54</td>
<td>8.16</td>
<td>9.38</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Receptive &amp; productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input only</td>
<td>24</td>
<td>8.79</td>
<td>1.93</td>
<td>7.98</td>
<td>9.61</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>9.19</td>
<td>1.69</td>
<td>8.42</td>
<td>9.96</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>8.52</td>
<td>2.00</td>
<td>7.69</td>
<td>9.35</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>70</td>
<td>8.81</td>
<td>1.88</td>
<td>8.37</td>
<td>9.26</td>
<td>4</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

This analysis yielded a significant effect of time, $F(1, 67) = 7.475, p = .008, \eta^2_p = .100$. Performance on all 24 items showed a statistically significant but small increase from 16.31 in the pretest to 17.59 in the posttest, that is, an increase of just over one point on a 24-point scale. No main effects of treatment or modality were obtained. In addition, the interactions were not significant.
Figure 4.10. Performance on receptive pretest and posttest by treatment group (maximum scores of 12)

Figure 4.11. Performance on productive pretest and posttest by treatment group (maximum scores of 12)
Subsequently, we conducted a three-way 3 (Treatment) x 2 (Time) x 4 (Outcome Measure) analysis. Numeric and graphic representations of the descriptives may be found in Table 4.4 and Figures 4.13 and 4.14.

Apart from the main effect of time (the same effect as in the previous analysis), $F(1, 67) = 7.475, p = .008, \eta^2_p = .100$, a main effect of outcome measure was obtained, $F(3, 201) = 3.117, p = .027, \eta^2_p = .044$. No significant interactions were obtained although the Treatment x Outcome Measure interaction approached significance, $F(6, 201) = 1.840, p = .063, \eta^2_p = .052$. Thus, the four subtests were not equally difficult for all groups but since the mean scores differ only marginally across treatment, time and outcome measure, we attribute no importance to the main effect of outcome measure.
Table 4.4. Performance on pretest and posttest items by outcome measure and by treatment group
(maximum scores of 6 on every outcome measure and 24 on all pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>Treatment</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI_{95%}</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR</td>
<td>GJ</td>
<td>Input practice</td>
<td>24</td>
<td>4.17</td>
<td>1.20</td>
<td>3.66</td>
<td>4.68</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>21</td>
<td>4.10</td>
<td>0.94</td>
<td>3.67</td>
<td>4.52</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control group</td>
<td>25</td>
<td>3.92</td>
<td>1.22</td>
<td>3.42</td>
<td>4.42</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>70</td>
<td>4.06</td>
<td>1.13</td>
<td>3.79</td>
<td>4.33</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Input practice</td>
<td>24</td>
<td>4.75</td>
<td>1.19</td>
<td>4.25</td>
<td>5.25</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>21</td>
<td>3.90</td>
<td>1.30</td>
<td>3.31</td>
<td>4.50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control group</td>
<td>25</td>
<td>4.28</td>
<td>1.31</td>
<td>3.74</td>
<td>4.82</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>70</td>
<td>4.33</td>
<td>1.29</td>
<td>4.02</td>
<td>4.64</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CCR</td>
<td>Input practice</td>
<td>24</td>
<td>4.33</td>
<td>1.09</td>
<td>3.87</td>
<td>4.79</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>21</td>
<td>3.86</td>
<td>1.39</td>
<td>3.23</td>
<td>4.49</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control group</td>
<td>25</td>
<td>4.28</td>
<td>0.98</td>
<td>3.88</td>
<td>4.68</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>70</td>
<td>4.17</td>
<td>1.15</td>
<td>3.90</td>
<td>4.45</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>TR</td>
<td>Input practice</td>
<td>24</td>
<td>3.42</td>
<td>1.02</td>
<td>2.99</td>
<td>3.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>21</td>
<td>3.67</td>
<td>1.32</td>
<td>3.07</td>
<td>4.27</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control group</td>
<td>25</td>
<td>4.16</td>
<td>0.90</td>
<td>3.79</td>
<td>4.53</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>70</td>
<td>3.76</td>
<td>1.11</td>
<td>3.49</td>
<td>4.02</td>
<td>1</td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.

Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Table 4.4. Performance on pretest and posttest items by outcome measure and by treatment group
(maximum scores of 6 on every outcome measure and 24 on all pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI&lt;sub&gt;low&lt;/sub&gt;</th>
<th>CI&lt;sub&gt;high&lt;/sub&gt;</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
<td>UL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSTTEST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>24</td>
<td>4.17</td>
<td>1.34</td>
<td>3.60</td>
<td>4.73</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Output practice</td>
<td>21</td>
<td>4.57</td>
<td>1.25</td>
<td>4.00</td>
<td>5.14</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Control group</td>
<td>25</td>
<td>4.24</td>
<td>1.45</td>
<td>3.64</td>
<td>4.84</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>70</td>
<td>4.31</td>
<td>1.35</td>
<td>3.99</td>
<td>4.64</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>24</td>
<td>4.67</td>
<td>1.44</td>
<td>4.06</td>
<td>5.27</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Output practice</td>
<td>21</td>
<td>4.29</td>
<td>1.88</td>
<td>3.43</td>
<td>5.14</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Control group</td>
<td>25</td>
<td>4.40</td>
<td>1.38</td>
<td>3.83</td>
<td>4.97</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>70</td>
<td>4.46</td>
<td>1.55</td>
<td>4.09</td>
<td>4.83</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CCR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>24</td>
<td>4.38</td>
<td>1.10</td>
<td>3.91</td>
<td>4.84</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Output practice</td>
<td>21</td>
<td>4.95</td>
<td>0.81</td>
<td>4.59</td>
<td>5.32</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Control group</td>
<td>25</td>
<td>4.24</td>
<td>1.05</td>
<td>3.81</td>
<td>4.67</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>70</td>
<td>4.50</td>
<td>1.03</td>
<td>4.25</td>
<td>4.75</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>24</td>
<td>4.42</td>
<td>1.50</td>
<td>3.78</td>
<td>5.05</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Output practice</td>
<td>21</td>
<td>4.24</td>
<td>1.18</td>
<td>3.70</td>
<td>4.77</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Control group</td>
<td>25</td>
<td>4.28</td>
<td>1.40</td>
<td>3.70</td>
<td>4.86</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>70</td>
<td>4.31</td>
<td>1.36</td>
<td>3.99</td>
<td>4.64</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.
Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Figure 4.13. Pretest and posttest performance on all four outcome measures by treatment group (maximum scores of 6 with GJ0 = grammaticality judgement pretest, GJ1 = grammaticality judgement posttest, etc.)
Given the results of the two analyses presented above, we did not deem it worthwhile to conduct more detailed analyses.

4.4 Conclusions and discussion

With respect to the three research questions formulated at the beginning of this chapter, the following conclusions can be drawn. Analyses of the data in Study 1 revealed a small increase from pretest to posttest performance (by 1.28 point, that is, from 16.31 to 17.59 on a total score of 24) but this effect was not mediated by treatment, modality or outcome measure. In other words, all three treatment groups slightly improved their performance from pretest to posttest. It is impossible to say that this improvement was caused by the three treatments since we had decided in advance not to integrate a no-treatment (control) group into the research design. We had expected to find a treatment effect with possibly the input-practice and output-practice groups performing better than the input-only group on the posttest. However, this expectation was not confirmed. Anticipating such an effect and considering the limited number of students available as participants, we had decided in advance to refrain from including a no-treatment control group in this study. Since there are no data from a no-treatment control group, it is impossible to say whether the small performance increases obtained in all three treatment groups were caused by the fact that participants in all three groups had received any treatment which dealt with the target structures under investigation. For a more detailed discussion and interpretation of the findings across all three of the studies, see Chapter 7, Sections 7.2 and 7.3.
We refrain from interpreting the small effect of outcome measure in the second analysis and the interaction effect which approached significance in the second analysis.

4.5 Implications for quantitative data collection in Study 2
In light of the results found after statistical analysis of the quantitative data in Study 1, several changes were made to the overall research design and research methodology for Study 2. Although these changes did not alter the general experimental setup beyond recognition, they were relevant in that they paved the way for—what was considered to be—an improved second study with a sharper focus on the relevant issues under investigation. All in all, a total of four changes were introduced for Study 2. A short description of these changes will be provided below. More detailed information about these changes and possible consequences for the overall research design and research methodology may be found in Chapter 5, which provides a detailed account of Study 2.

The first change which was implemented was a change in the number of treatments, which was reduced from three to two. Since the experimental focus was on possible differences between receptive and productive treatments in a specific pedagogical setting, in which explicit FFI was the norm (i.e., explicit FFI in a translation programme), the treatment which consisted of mere input enhancement (termed input only in Study 1) did not truly meet the requirements of explicit FFI which could possibly be administered to the participants in reality. Instead, the two remaining treatments—input practice and output practice—were retained for further investigation in Study 2 since they were both feasible operationalizations of explicit FFI instruction.

The second change was one related to the type of practice administered in the input-practice and output-practice groups. Taking into account the outcome measures being used in the studies, the participants’ prior experience with explicit FFI instruction on the English tense system and the participants’ academic environment at the time the research was being carried out, we decided to gear the practice towards the four outcome measures in use for the study. In practice, this meant that the input-practice group would receive practice for grammaticality judgement task types and selected response task types. The output-practice group, on the other hand, would receive practice for constrained constructed response task types and translation task types.

A third change was related to the number of test items. In Study 1, a total of 24 items were used in the pretest and posttest. These items were evenly distributed across the four outcome measures, resulting in six items per outcome measure. For Study 2, it was decided that eight additional items should be added, bringing the total to 32 test items, with eight test items per outcome measure. The addition of eight new items was deemed feasible within the time constraints of the sessions being carried out. The increase in test items also provided additional data to work with when analysing the data in Study 2.
The fourth and final change was related to the third change described above. Adding items to the existing test format implies either longer testing times or increased time pressure for the participants. Since neither was desirable, it was decided that the participants should receive only tense-related questions during the pretest and posttest sessions. Instead of asking them to answer one tense-related question per test in addition to three questions which were not related to tense, the only questions that they were asked were tense-related question, with one question per text. An added advantage of limiting the questions to tense-related questions was that the pretest and posttest sessions no longer had to be spread over two weeks. The participants were able to complete each of these sessions in one week.

### 4.6 Qualitative results

For Study 1, the quantitative data reported above were the most substantial part of the data sets that were collected. However, a decision was taken to collect some qualitative data too in an attempt to gain more comprehensive insights into the strategies with which the participants approached the pretest and posttest task types. Not only did we want to see what the participants answered, we also wanted the participants to shed light on how they came to their answers. Although the participants’ answers to the tense-related questions were vital to the qualitative analysis of the data found in Study 1, they were not the focus of investigation. The focus was more on the strategies that the participants used to answer the tense-related questions.

A total of 24 participants were invited to take part in think-aloud protocols (TAPs). During the TAPs, the participants were presented with twelve texts (three for every outcome measure), which they had already received during the experimental pretest and/or posttest sessions. The TAP participants were selected based on their pretest and posttest scores during the quantitative data collection procedures, with eight participants being selected from the following three scoring scales: ≤50%, 51-75%, 76-85%.

The TAP participants were asked to provide answers to the text-related questions, which they had already been asked during the pretest and posttest experimental sessions. The questions that the TAP participants were asked were identical with the questions that they had been asked during the experimental pretest and posttest sessions. However, this time the questions were asked orally and the TAP participants were asked to answer the questions whilst simultaneously verbalizing their thoughts, strategies, choices and final answers, which were all recorded for subsequent qualitative analysis.

As far as their answers to the tense-related questions were concerned, the TAP participants showed overall learner variability, which proved to be both positive and negative variability. The participants’ answers recorded during the TAPs (October 2008) were compared with their answers (to identical tense-related questions) recorded during the posttest (May 2008). Participants who
provided grammatically correct answers during the posttest and grammatically incorrect answers (on the same tense-related questions) during the TAPs were said to show negative variability. Participants who provided grammatically incorrect answers during the posttest and grammatically correct answers (on the same tense-related questions) during the TAPs were said to show positive variability. A more detailed look at the answers and, more important, the strategies that the participants employed to answer the tense-related questions revealed the following results.

The TAP participants often relied on intuition or, as they described it, ‘feeling’ to answer tense-related feelings. Answers such as “It feels right” or “It sounds better” were often provided. When the TAP participants were asked to provide additional information about such subjective explanations, they were often unable to do so. Many of the TAP participants had to be nudged into verbalizing their thoughts, strategies, choices and final answers. Overall, the answers to tense-related pictures ranged from correct answers to incorrect answers with nuanced variations on these two options depending on whether the answers and/or the metalinguistic explanations were right. For example, some of the TAP participants were able to provide both the grammatically correct answers to tense-related questions (+) and the correct metalinguistic explanations about their answers (+). By way of contrast, some TAP participants provided grammatically incorrect answers to tense-related questions (−) and were also unable to provide any correct metalinguistic explanations (−). However, there were also TAP participants who provided grammatically correct answers to tense-related questions (+) but were unable to provide any correct metalinguistic explanations (−). In other words, these participants showed (signs of) grammatically correct ESL performance, which was not necessarily based on correct verbalized metalinguistic explanations. In addition, a fourth group of TAP participants was recognized: Some TAP participants were unable to provide grammatically correct answers to tense-related questions (−) even though they provided correct metalinguistic explanations (+).

Many of the TAP participants showed strategies which indicated an awareness of temporal discourse features. Not only did the TAP participants take into account the use of, for example, adverbials, they also appeared receptive to applying pragmatic (world) knowledge when deciding which tense to use. In addition to these two features of temporality, the TAP participants also used scaffolding features to decide on tense usage. Often, they expressed awareness of tenses already in use in either the same sentence in which the targeted verb form appeared or in more extended discourse. However, awareness of scaffolding features was not necessarily a guarantee for answering the tense-related questions correctly. By way of conclusion, it can be said that even though many

96 A more nuanced picture revealed that some TAP participants were able to provide snippets of correct metalinguistic explanations but unable to provide complete explanations.
of the TAP participants showed an awareness of temporal features, this awareness was not always strongly developed that it led to consistently correct answers to tense-related questions.

There were also TAP participants who applied L1-based strategies, stating that they firmly believed that English and their L1 (i.e., Dutch) often behaved in exactly the same way and that such likeness between the two languages was the overall strategy which guided them when selecting possible tense forms in English.

When asked about the overall difficulty related to the English tense system, most of the TAP participants agreed that the English tense system—whatever this entailed for them in detail—was quite problematic for them. They expressed doubts about their theoretical and practical knowledge about the English tense system. Although they generally agreed that form was not a problem per se, they did find the meanings and the uses of the tenses complex at times. When asked about the difficulty of the four task types that they had encountered during the experimental pretest and posttest sessions, the TAP participants were not all in agreement. Many found the translation task type the most difficult since they felt uncomfortable about translating from Dutch into English without any dictionaries to look up words and expressions which they regarded as difficult. There were, however, also a number of TAP participants who expressed their unease about the two receptive task types (grammaticality judgement, selected response). They admitted to having doubts when not having to produce verb forms themselves and being given tensed verb forms to interact with.

For some of the TAP participants, general doubts were so deeply rooted that they felt that they had to “think like a teacher”. In practice, this meant that if they were given a text and were asked about a finite verb form, they could not believe that the teacher would select, for example, a past tense as the correct finite verb form for the third time in a row. So the strategy that they then applied was that the finite verb form had to be a present perfect because of the two previous finite verb forms, which had already been past forms.
CHAPTER 5
STUDY 2

A statistical analysis, properly conducted, is a delicate dissection of uncertainties, a surgery of suppositions
(M. J. Moroney)

5.1 Introduction

The insights gained from the data analyses for Study 1 were used as a starting point for the design for Study 2. As discussed in Chapter 4, Section 4.5, four changes were implemented in the setup of Study 2. By way of summary, these four changes are listed below:

1. A reduction of the number of treatments from three to two, resulting in an input-practice group and an output-practice group (and the omission of the input-only group),
2. A focus on pretest and posttest outcome measures during the practice sessions, resulting in grammaticality judgement (GJ) and selected response (SR) practice for the input-practice group, and constrained constructed response (CCR) and translation (TR) practice for the output-practice group,
3. An evenly distributed increase of the pretest and posttest items from 24 to 32 over the four outcome measures as a whole, resulting in two additional items per outcome measure,
4. An exclusive focus on tense-related questions during the pretest and posttest sessions.

Since many of the design-related and methodological features for Study 2 were identical with those found in Study 1, I will refrain from consistently pointing out the details related to any overlap between both studies. In the event of any large overlaps, the reader will be advised to consult the information described in detail for Study 1 in Chapter 4. For example, the focus in the Methods section in this chapter (Section 5.2) will be on the changes that were implemented for Study 2. Overlaps between Studies 1 and 2 will occasionally be highlighted but unless differences are explicitly discussed, the reader should view the setup of Study 1 as identical with the setup of Study 2.

Once again, the general research question for Study 2 was whether explicit FFI had any effect on the acquisition of the complex past/present perfect distinction (when used to locate bygone situations) by Dutch-speaking ESL learn-
ers. For Study 2, the three research questions for Study 1 were formulated once again and are provided below by way of recapitulation.

**Research questions**

1. Does explicit form-focused instruction (FFI) have an effect on (Dutch-speaking) ESL learners’ performance with regard to complex temporal FMU mappings (i.e., the past/present perfect distinction) in L2 English?

2. If explicit FFI does have an effect, can any overall differential effects be ascertained with respect to the specific type of instructional treatment (e.g., input practice, output practice)?

3. If differences between treatments can be ascertained, are the differential effects the same across all of the outcome measures? Or are there differences between receptive and productive outcome measures?

The research hypotheses accompanying the three research questions were identical with the ones formulated for Study 1. Once again, we did not include a no-treatment (control) group to address research question 1. We simply assumed that an effect of treatment would be visible on posttest performances.

**5.2 Method**

**5.2.1 Design**

For Study 2, the four-way mixed-design template which was used for Study 1 was re-used as a design template. However, one structural change was made. Instead of assigning three levels to the between-subjects variable treatment—as was the case in Study 1 (input only, input practice, output practice)—we assigned two levels to the treatment variable in Study 2 (input practice, output practice). In other words, we did away with the input-only level. The within-subjects variables were once again time, with two levels (pretest, posttest), modality, with two levels (receptive, productive) and outcome measure, with four levels (grammaticality judgement (GJ), selected response (SR), constrained constructed response (CCR) and translation (TR)). Consequently, Study 2 was drawn up using a four-way 2 (Treatment) x 2 (Time) x 2 (Modality) x 4 (Outcome Measure) mixed-design template. The within-subjects variables were nested in exactly the same way as in Study 1, with modality nested hierarchically under time, with outcome measure nested non-hierarchically under modality (see Figure 5.1). Consequently, it was once again inappropriate to conduct an analysis including all independent variables since the four outcome measures were split across the two modalities, with GJ and SR nested under the receptive level of the modality variable and CCR and TR nested under the productive level of the modality variable. As a result, we decided once again to carry out two separate analyses (see Section 5.3).
dependent variable

1

continuous

independent variables

4
categorical

1
between
groups

3
within
groups

written performance on a 32-item tense-related test consisting of four separate 8-item outcome measures of either a receptive or a productive nature: (1) grammaticality judgement, (2) selected response, (3) constrained constructed response and (4) translation

Between-groups variable:
1) Treatment (2 levels)
   1) Input practice
   2) Output practice

Within-groups variables:
1) Time (2 levels)
   1) Pretest
   2) Posttest
2) Modality (2 levels)
   1) Receptive
   2) Productive
3) Outcome Measure (4 levels)
   1) Grammaticality judgement
   2) Selected response
   3) Constrained constructed response
   4) Translation

Figure 5.1. Design box for Study 2 (2 x 2 x 2 x 4 mixed design)

5.2.2 Participants
The same population of ESL learners as the one under investigation in Study 1 was used for the selection of a sample of participants for Study 2. The only difference was that the participants were new and had not participated in Study 1. Because there were students who had not passed the English Grammar course the year before and, consequently, had already taken part in Study 1, we decided to allow these students to take part in Study 2 but their data were not used to avoid the influence of confounding variables (e.g., carry-over effects). Any data reported below automatically exclude participants who were retaking the English Grammar course and, consequently, had already participated in Study 1.
**Attrition.** Initially, 88 participants took part in the pretest but the number of participants dropped to 79 during the four experimental sessions. This drop corresponded with an attrition rate of approximately 10%. Once again, the drop in participant numbers was the result of extraneous factors (see Chapter 4, Section 4.2.2, Attrition). For the analysis and discussion of the experimental data, the 10% attrition rate was taken into consideration. We also tried to avoid any ceiling effect from influencing the data by screening the pretest data and excluding—as we did for Study 1—any participants who scored more than 85% of the pretest items correctly. In practice, 85% on the pretest corresponded with a score of 28 or higher on a possible maximum score of 32. The data from eight participants were not taken into consideration as a result of this screening procedure. After taking into consideration the attrition rate and the screening procedure, 71 participants remained. The data from these 71 participants were considered in the descriptive and inferential data analyses.

**Assignment to treatment groups.** The participants were assigned to treatment groups in the same way as in Study 1. This resulted in random assignment to one of the two treatment groups as follows: 36 (input practice) and 35 (output practice).

**ESL grammar course.** The experimental sessions were once again conducted as part of the participants’ regular English Grammar course in the first year of their English bachelor programme. The only difference was that Study 2 took place in the academic year 2008-2009, which was a year after Study 1. All the other course-related information was the same as the information provided for Study 1 (see Chapter 4, Section 4.2.2, ESL grammar course).

**ESL language proficiency.** In line with the criteria used for Study 1, the participants’ overall level of proficiency for English was categorized as intermediate to upper intermediate. The categorization criteria which were used for Study 2 were the same as those used for Study 1.

**Academic exposure to ESL English.** As far as the academic exposure to ESL English was concerned, no changes were reported with the situation described for Study 1.

**Academic exposure to other L2s.** The foreign language combinations were once again dependent on the participants’ own choices within the constraints of the bachelor programme structure in use. The context in which the foreign language combinations were made was described in Chapter 4 so the reader is advised to consult the details reported there for any background information regarding the participants’ foreign language combinations (Chapter 4, Section 4.2.2, Academic exposure to other L2s). The descriptives for the 71 participants’ foreign language combinations found in Study 2 are presented in Table 5.1.
Of the 71 participants, 65 had selected English as part of an L2 combination consisting of two L2s. As far as instruction was concerned, these 65 participants received seven hours of ESL instruction on a weekly basis. For their other L2s in their L2 combinations, they also received seven hours of instruction on a weekly basis. As was the case for all of the other first-year participants, these 65 participants were expected to take Dutch language courses since Dutch was their A language. The remaining six participants had chosen English as part of an L2 combination consisting of three L2s. These participants received either six or seven hours of English on a weekly basis, which they combined with two other L2s and Dutch (as their A language).

**Gender and age.** Of the 71 participants, 21 were male and 50 were female. The participants ranged in age from 18 to 29 years and the mean age was 20 years ($SD = 2$, mode = 19, median = 19).

**L1 backgrounds.** As was the case for the participants in Study 1, the participants in Study 2 were asked about their L1 backgrounds and about which language(s) that they themselves regarded as their L1(s). Of the 71 participants, 70 indicated that they had one L1. Of those 70 participants, 68 participants had Dutch as their L1s. The remaining two participants’ L1s were Albanian and Polish. One participant indicated that she had two L1s, which were Dutch and Russian.

**Remuneration.** The same remuneration information as that provided for Study 1 was applied for Study 2.
5.2.3 Apparatus and materials

5.2.3.1 Apparatus
The participants completed all of the four experimental sessions in the same computer language lab that was used for Study 1. In addition, the infrastructure in the language lab was the same as that found in Study 1. The only difference were the updated computers and computer programs which the participants used for the experimental sessions. Since these updates were of no significance for the actual execution of the individual experimental sessions, no additional update-related information will be provided.

For Study 2, the computer specialist who was contracted for Study 1 was once again asked to oversee the programming of the software. Apart from some design-related changes (see Section 5.2.1), no fundamental changes were made to the programming of the experimental sessions. Technical support was again available throughout the entire duration of the four experimental sessions and no major computer-related and/or network-related problems were reported.

5.2.3.2 Materials
For a detailed description of the materials used for Study 2, the reader is advised to read Chapter 4, Section 4.2.3.2, which provides in-depth information about the materials used. Since most of this information did not change for Study 2, the description provided in this section will be limited to the changes that were made to the materials. The changes will be discussed in the following subsections: pretest, treatment theory, treatment practice and posttest. This subdivision reflects the setup of the experimental sessions in Study 2.

Pretest. The corpus of texts which was used for Study 1 was also used for Study 2. For information related to the creation of the corpus, see Chapter 4, Section 4.2.3.2, Pretest. However, instead of selecting 24 items for the pretest, we decided to increase the number of pretest items to 32 items. This increase allowed us to gather more data and to do so across the four outcome measures, with two items being added to every outcome measure. A second change in the pretest was the number of questions per text. Whereas in Study 1 we had decided to present the participants with one tense-related question and three additional questions related to other aspects of the texts (e.g., cultural references, grammar items not related to tense, lexis), we decided to do away with the questions which were not tense-related. The main advantage of asking only one (tense-related) question per text was that the pretest became less time-consuming and could be taken in one experimental session, which was beneficial for the internal validity of Study 2. In Study 1, the participants were given two experimental sessions to complete the pretest because they were asked to answer four questions per test item.

The types of tense-related questions that were asked were identical with the ones that were asked in the pretest for Study 1. For additional information on
the nature of the tense-related questions and for examples of such questions, see Chapter 4, Section 4.2.3.2, Pretest.

**Treatment theory.** No changes were made to the treatment theory. All of the participants received the exact same theoretical information. For detailed information on the theory session, see Chapter 4, Section 4.2.3.2, Treatment theory.

**Treatment practice.** For Study 2, there were two experimental groups—input practice and output practice—which received different forms of tense-related treatment practice. Although the two experimental groups were given names which are identical with two of the three treatments in Study 1, the nature of the forms of treatment practice in Study 2 were structurally different from the forms of treatment practice in Study 1. The seven texts that were used for the treatment practice in Study 1 were re-used (with the addition of one text to bring the number of texts to eight, which allowed for an even distribution of texts across the four outcome measures) but the nature of the practice was changed. In other words, the treatment practice reflected more closely the input/output distinction found for the pretest and posttest outcome measures in Study 2. In essence, the practice was geared towards the four outcome measures used for Study 2. This meant that the input-practice group received practice for grammaticality judgement (GJ) task types and selected response (SR) task types whereas the output-practice group received practice for constrained constructed response (CCR) task types and translation (TR) task types. The materials that were used for the practice sessions consisted of eight texts, which were identical for both instructional treatments. For the practice geared towards the GJ task types, the verb form of interest (i.e., either a past or a present perfect) changed colour (from black to blue) and appeared in bold in the text 40 seconds after the text had appeared on-screen. Subsequently, the input-practice participants received a question about the verb form in bold. The input-practice participants were then requested to select one of the two options (see Figure 5.2).

Upon confirmation of their answers, the input-practice participants received a slide containing metalinguistic feedback on the linguistic context in which both the finite verb form and other relevant temporal information available in the text were highlighted. The feedback provided to the input-practice participants in Study 2 was identical with the feedback provided to the input-practice participants in Study 1. For the treatment practice geared towards the SR task types, the input-practice participants received a text with a selected response task type (see Chapter 4, Figure 4.3). This was followed by a picture selection task with feedback. The output-practice participants received the same texts as the input-practice participants but with CCR task types and TR task types (see Chapter 4, Figures 4.4 and 4.5). Additionally, the output-practice participants also received the picture selection task with feedback, which the input participants also received.
Louis Armstrong (1901-1971) was one of the greatest jazz trumpet players of the 20th century, but he ended his career though as a singer. Lips are not suited to vibrate against the mouthpiece of a trumpet day in and day out. Armstrong struggled with this fact all his life. Even at the height of his career in the 1930s, his lips had been ruined as a result of his energetic blowing technique. He was always treating his injuries and blisters using cream. Every time he put his instrument against his lips it felt as though a scorching poker was being held against them. At the end of a concert he was forced to swallow the constant stream of blood. There was a concert in 1935 during which the blood flowed down his shirt during his performance. In the end, Armstrong was forced to cancel so many performances that his manager sued him for breach of contract.

**Question**

In standard, formal present-day English, the verb form in bold in the text above is considered . . .

- **grammatically correct**
- **grammatically incorrect**

**Figure 5.2. Example text and question for input-practice group**

**Posttest.** One week after completing the treatment sessions, the participants were asked to take part in an unannounced posttest. The setup of the posttest in Study 2 was identical with the setup of the pretest. Once again, the difference was the actual nature of the posttest items. Whereas the pretest consisted of 32 new items, the posttest consisted of 16 ‘old’ items (i.e., items which had been recycled from the four pretest outcome measures, with four recycled items per outcome measure) and 16 ‘new’ items, which were new for all of the participants.

**5.2.4 Procedure**

Study 2 was organized in the month of March of the academic year 2008-2009. It consisted of four experimental sessions, which were divided as follows: pretest (one session), treatment theory (one session), treatment practice (one session), posttest (one session).

Once again, the participants were invited to enrol before the beginning of the first experimental session. The participants were told that the experimental sessions would focus on an aspect of English grammar but no further information was provided to prevent the participants from preparing for the experimental sessions.
During the first experimental session, the participants were informed that they would be taking a test and that any test-related information would be provided to them on-screen. The instructions that the participants received for the pretest in Study 2 were almost identical with the instructions given to the participants in Study 1. The only difference was that the participants were instructed to complete the entire pretest in one experimental session. One week later, the participants received explicit information about the ESL target features under investigation. A week after receiving this information the participants took part in a practice session, which was different depending on which of the two treatments the participants had been assigned to. Since there were no changes for these two sessions between Studies 1 and 2, no further procedural information will be provided here. A week after completing the practice session, the participants were invited to take a posttest. Like the pretest, the posttest consisted of 32 items. The selection of the 32 posttest items for Study 2 was completed in the same way as the selection of the 24 posttest items for Study 1 (see Chapter 4, Section 4.2.4).

5.2.5 Coding and scoring
No changes were made to the coding and scoring procedure. In other words, the procedures explained for Study 1 in Chapter 4, Section 4.2.5, were replicated and used for Study 2.

5.2.6 Statistical analyses
With regard to the statistical analyses, the same logic that was applied to the design in Study 1 was applied to the design in Study 2. Instead of running a 2 (Treatment) x 2 (Time) x 2 (Modality) x 4 (Outcome Measure) analysis, we decided to run two separate analyses (see Section 5.3). For any other details concerning the statistical analyses, the reader is advised to read Chapter 4, Section 4.2.6.

5.2.7 Validity and reliability
Since we used the same corpus of texts for the selection of pretest and posttest items, the information provided in Chapter 4, Section 4.2.7, on the validity of the items also applies to the texts which were used for Study 2. Reliability was once again measured using Cronbach’s alpha coefficient (α). The Cronbach’s alpha coefficients for the pretest and posttest in Study 2 were .63 and .73 respectively.
5.3 Results

No qualitative data were collected for Study 2. Consequently, the results reported in this section are all based on the quantitative data, which were collected throughout the pretest and posttest sessions.

Once again, analyses for both main effects and interaction effects were carried out at all of the three levels of the experimental design as indicated in Figure 4.9 in Chapter 4.

As mentioned in Section 5.2.2, eight students scored 28 or higher on the pretest. In other words, they answered more than 85 per cent of the pretest items correctly. As was the case for Study 1, we took these students to be familiar with the ESL target structures under investigation and instruction was not likely to improve their performance significantly. Consequently, we conducted the two analyses on only those participants with pretest scores of 27 or lower, leaving 36 and 35 participants in the input-practice and output-practice groups respectively. We first conducted a three-way 2 (Treatment) x 2 (Time) x 2 (Modality) RM ANOVA, with treatment as the between-subjects variable and time and modality as the within-subjects variables. The descriptives are shown numerically in Table 5.2 and graphically in Figures 5.3, 5.4 and 5.5.
CHAPTER 5

Table 5.2. Performance on receptive and productive pretest and posttest items by treatment group (maximum scores of 16 on receptive and productive items and 32 on the totality of pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Modality</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI_{LL}</th>
<th>CI_{UL}</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Receptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>10.14</td>
<td>2.24</td>
<td>9.38</td>
<td>10.90</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>10.00</td>
<td>2.80</td>
<td>9.04</td>
<td>10.96</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>10.07</td>
<td>2.52</td>
<td>9.48</td>
<td>10.67</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>10.39</td>
<td>2.64</td>
<td>9.49</td>
<td>11.28</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>10.11</td>
<td>1.88</td>
<td>9.47</td>
<td>10.76</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>10.25</td>
<td>2.29</td>
<td>9.71</td>
<td>10.79</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Receptive &amp; productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>20.53</td>
<td>3.92</td>
<td>19.20</td>
<td>21.85</td>
<td>13</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>20.11</td>
<td>4.19</td>
<td>18.68</td>
<td>21.55</td>
<td>13</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>20.32</td>
<td>4.03</td>
<td>19.37</td>
<td>21.28</td>
<td>13</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Receptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>13.14</td>
<td>2.05</td>
<td>12.45</td>
<td>13.83</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>12.49</td>
<td>2.48</td>
<td>11.63</td>
<td>13.34</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>12.82</td>
<td>2.28</td>
<td>12.28</td>
<td>13.36</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>12.03</td>
<td>2.46</td>
<td>11.20</td>
<td>12.86</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>11.74</td>
<td>2.44</td>
<td>10.90</td>
<td>12.58</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>11.89</td>
<td>2.44</td>
<td>11.31</td>
<td>12.46</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Receptive &amp; productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>36</td>
<td>25.17</td>
<td>3.78</td>
<td>23.89</td>
<td>26.45</td>
<td>16</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>24.23</td>
<td>4.39</td>
<td>22.72</td>
<td>25.74</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>24.70</td>
<td>4.09</td>
<td>23.74</td>
<td>25.67</td>
<td>13</td>
<td>30</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

This analysis yielded a significant effect of time, $F(1, 69) = 131.060, p = .000$, $\eta^2_p = .655$. Performance on all 32 items showed a statistically significant, large increase from 20.32 in the pretest to 24.70 in the posttest, that is, an increase of slightly more than four points on a 32-point scale. No main effects of treatment or modality were obtained. In addition, most of the interactions were not significant. The only interaction which was significant was the Time x Modality interaction, $F(1, 69) = 9.319, p = .003, \eta^2_p = .119$ (see Figure 5.6). However, since the mean scores differ only marginally across time and modality, we attribute no importance to this finding in light of the research questions under investigation.
Figure 5.3. Performance on receptive pretest and posttest by treatment group (maximum scores of 16)

Figure 5.4. Performance on productive pretest and posttest by treatment group (maximum scores of 16)
Subsequently, we conducted a three-way 2 (Treatment) x 2 (Time) x 4 (Outcome Measure) analysis. Numeric and graphic representations of the descriptives may be found in Table 5.3 and Figures 5.7, 5.8 and 5.9.
Apart from the main effect of time (the same effect as in the previous analysis, \( F(1, 69) = 131.060, p = .000, \eta^2_p = .655 \)), a main effect of outcome measure was obtained, \( F(3, 207) = 13.356, p = .000, \eta^2_p = .162 \). There was one significant interaction, which was the Time x Outcome Measure interaction, \( F(3, 207) = 8.588, p = .000, \eta^2_p = .111 \). Both the descriptive and inferential statistics showed that the four outcome measures were not equally difficult but since the mean scores differ only marginally across treatment, time and outcome measure, we attribute no importance to the main of effect of outcome measure and the Time x Outcome Measure interaction in light of the research questions which were investigated.

Table 5.3. Performance on pretest and posttest items by outcome measure and by treatment group
(maximum scores of 8 on every outcome measure and 32 on all pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Treatment</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI95%</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRETEST</td>
<td>Input practice</td>
<td>36</td>
<td>4.72</td>
<td>1.21</td>
<td>4.31</td>
<td>5.13</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>4.34</td>
<td>1.63</td>
<td>3.78</td>
<td>4.90</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>4.54</td>
<td>1.43</td>
<td>4.20</td>
<td>4.87</td>
<td>1</td>
</tr>
<tr>
<td>CCR</td>
<td>Input practice</td>
<td>36</td>
<td>5.64</td>
<td>1.61</td>
<td>5.10</td>
<td>6.18</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>5.51</td>
<td>1.34</td>
<td>5.06</td>
<td>5.97</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>5.58</td>
<td>1.47</td>
<td>5.23</td>
<td>5.93</td>
<td>1</td>
</tr>
<tr>
<td>TR</td>
<td>Input practice</td>
<td>36</td>
<td>4.75</td>
<td>1.46</td>
<td>4.26</td>
<td>5.24</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>35</td>
<td>4.60</td>
<td>1.19</td>
<td>4.19</td>
<td>5.01</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>71</td>
<td>4.68</td>
<td>1.33</td>
<td>4.36</td>
<td>4.99</td>
<td>2</td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.
Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Table 5.3. (Contd.). Performance on pretest and posttest items by outcome measure and by treatment group  
(maximum scores of 8 on every outcome measure and 32 on all pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>Treatment</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>POSTTEST</td>
<td>GJ</td>
<td>Input practice</td>
<td>36</td>
<td>6.42</td>
<td>1.27</td>
<td>5.99</td>
<td>6.85</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>GJ</td>
<td>Output practice</td>
<td>35</td>
<td>6.29</td>
<td>1.25</td>
<td>5.86</td>
<td>6.72</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>GJ</td>
<td>Mean</td>
<td>71</td>
<td>6.35</td>
<td>1.26</td>
<td>6.06</td>
<td>6.65</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Input practice</td>
<td>36</td>
<td>6.72</td>
<td>1.19</td>
<td>6.32</td>
<td>7.12</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Output practice</td>
<td>35</td>
<td>6.20</td>
<td>1.64</td>
<td>5.64</td>
<td>6.76</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Mean</td>
<td>71</td>
<td>6.46</td>
<td>1.44</td>
<td>6.12</td>
<td>6.81</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>CCR</td>
<td>Input practice</td>
<td>36</td>
<td>6.31</td>
<td>1.19</td>
<td>5.90</td>
<td>6.71</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>CCR</td>
<td>Output practice</td>
<td>35</td>
<td>6.23</td>
<td>1.42</td>
<td>5.74</td>
<td>6.71</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>CCR</td>
<td>Mean</td>
<td>71</td>
<td>6.27</td>
<td>1.30</td>
<td>5.96</td>
<td>6.57</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>TR</td>
<td>Input practice</td>
<td>36</td>
<td>5.72</td>
<td>1.91</td>
<td>5.08</td>
<td>6.37</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>TR</td>
<td>Output practice</td>
<td>35</td>
<td>5.51</td>
<td>1.48</td>
<td>5.00</td>
<td>6.02</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>TR</td>
<td>Mean</td>
<td>71</td>
<td>5.62</td>
<td>1.70</td>
<td>5.22</td>
<td>6.02</td>
<td>1</td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.
Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Figure 5.7. Pretest and posttest performance on all four outcome measures by treatment group (maximum scores of 8 with GJ0 = grammaticality judgement pretest, GJ1 = grammaticality judgement posttest, etc.)
Figure 5.8. Pretest and posttest mean performance by treatment group (maximum scores of 8)

Figure 5.9. Interaction between overall mean performance on outcome measures and time of testing (maximum scores of 8)
In light of the results of the two analyses presented above, we did not deem it worthwhile to pursue any subsequent analyses.

5.4 Conclusions and discussion

With respect to the three research questions formulated at the beginning of this chapter, the following conclusions can be drawn. Analyses of the data in Study 2 revealed a large increase from pretest to posttest performance. This increase was a 4.38-point increase when comparing overall pretest and posttest scores (20.32 and 24.70 respectively on a total score of 32). However, it was impossible to state that this effect was mediated by treatment, modality or outcome measure. In other words, the two treatment groups improved their performance from pretest to posttest but that is all that can be said.

We had expected to find a treatment effect. However, anticipating such an effect and considering the limited number of students available as participants, we had decided in advance to refrain once again from including a no-treatment (control) group in Study 2. Since no data from a no-treatment control group are available, it is impossible to say whether the relatively large performance increases obtained were caused by the fact that participants in both groups had received any treatment which dealt with the target structures under investigation.

Although we found additional main effects and interaction effects, these effects were moderate and did not shed any light on the research questions that were drawn up for the study. Consequently, we decided to refrain from interpreting these effects. For a more detailed discussion and interpretation of the findings across all three of the studies, see Chapter 7, Sections 7.2 and 7.3.

5.5 Implications for quantitative data collection in Study 3

In light of the results found after statistical analysis of the quantitative data from Study 2, we decided to make only one fundamental change to the overall research design and research methodology for a third study. The change that was implemented was the inclusion of a no-treatment group. In all other respects, Study 3 was an exact replication of Study 2. The addition of a no-treatment control group was deemed essential for interpreting the data and for drawing conclusions with respect to any possible effect of treatment. Studies 1 and 2 did not have any no-treatment groups so any significant effects that were found in those studies could not be based with any degree of certainty on the treatment that was provided. To allow for more certainty in the interpretation of possible effects, a no-treatment group was added for Study 3. Chapter 6 provides a detailed overview of Study 3.
CHAPTER 6
STUDY 3

Problems are not stop signs, they are guidelines.
(Robert Schuller)

6.1 Introduction
Study 3, which will be reported on in this chapter, is the last of the experimental studies that were carried out for this doctoral dissertation. In essence, it was a replication of Study 2 with two fundamental modifications. For Studies 1 and 2, we had expected in advance to find a treatment effect. Consequently, we had not addressed the first of three overarching research questions across all three studies. Although the results in Studies 1 and 2 appeared to confirm this expectation, we were unable to address the issue of treatment effect directly since we had consciously decided not to include a no-treatment (control) group in the research design and methodology. In an attempt to remedy this uncertainty about any treatment effect, we decided to set up Study 3 with the two treatment groups that were used for Study 2 (input practice, output practice) and to include this time a no-treatment control group. The inclusion of the no-treatment control group was the first of two modifications. In addition, we decided to introduce a change, which was set up to shed light on the durability of any possible treatment effects. The issue of treatment-effect durability was addressed by inviting the participants to participate in both an immediate posttest and a delayed posttest. In all other respects, Study 3 was identical with Study 2. Consequently, the reader will find many cross-references to the design-related and methodological features described in Chapters 4 and 5. I will once again refrain from consistently highlighting every overlap between Study 3 on the one hand and Studies 1 and 2 on the other hand. Any modifications, however, will be explained in detail.

The general research question for Study 3 was whether explicit FFI had any effect on the acquisition of the complex past/present perfect distinction (when used to locate bygone situations) by Dutch-speaking ESL learners. This led us to formulate the same research questions as the ones formulated in Studies 1 and 2. However, because of the two modifications that were introduced, we expanded the research focus to include two additional research questions. The following four research questions were formulated:
Research questions

1. Does explicit form-focused instruction (FFI) have an effect on (Dutch-speaking) ESL learners’ performance with regard to complex temporal form–meaning–use mappings (i.e., the past/present perfect distinction) in L2 English?
2. If explicit FFI does have an effect, can any overall differential effects be ascertained with respect to the specific type of instructional treatment (e.g., input practice, output practice)?
3. If differences between treatments can be ascertained, are the differential effects the same across all of the outcome measures? Or are there differences between receptive and productive outcome measures?
4. Are any treatment effects of a durable nature?

Research questions 2 and 3 were formulated and addressed in the setup for Studies 1 and 2 so I will assume that the reader is already familiar with them. By including this time a no-treatment (control) group, we were able to formulate and address research question 1, which looks at the issue of FFI having any possible effect on ESL learners’ performances. The inclusion of a delayed post-test to investigate the durability of any treatment effects was addressed by research question 4.

The research hypotheses that may be formulated for the research questions which were not addressed in Studies 1 and 2—research questions 1 and 4—are the following:

Research hypotheses

1. $H_0$: Explicit FFI does not have an effect on Dutch-speaking ESL learners’ performance with regard to the selected complex ESL target features.

$H_1$: Explicit FFI does have an effect on Dutch-speaking ESL learners’ performance with regard to the selected complex ESL target features.

4. $H_0$: The treatment effects are not of a durable nature.

$H_1$: The treatment effects are of a durable nature.

With respect to the hypotheses formulated for the first research question, it was expected that explicit FFI would indeed have an effect on the (Dutch-speaking) ESL learners’ test performances as far as the selected complex ESL target features were concerned. The English tense system is a staple in most ESL courses and many tense-related issues are relatively straightforward for Dutch-speaking ESL learners. However, there are some temporal features which cause prob-
lems for Dutch-speaking ESL learners regardless of their levels of proficiency (see Chapters 1 and 2). The targeted complex past/present perfect distinction is an example of such a problematic feature. An explicit focus on the (processing) problems that Dutch-speaking ESL learners experience by means of (contrastive) metalinguistic information about the ESL target features was expected to be beneficial when combined with receptive or productive forms of practice treatment.

As far as the fourth research question was concerned, it was hoped that instruction would have durable effects but considering the complexity of the target feature under investigation and factors related to the nature of the instruction (e.g., treatment duration), there was a chance that instructional effects would not be of a durable nature.

6.2 Method

6.2.1 Design

For Study 3, the four-way mixed-design template which was used for Studies 1 and 2 was put to use once again. By including a no-treatment group, we were, in effect, adding an additional level to the between-subjects variable treatment resulting in three levels for Study 3 (input practice, output practice, no treatment). The design-related features for Study 3 are represented in Figure 6.1. The within-subjects variables were once again time, this time, however, with three levels (pretest, immediate posttest, delayed posttest), modality, with two levels (receptive, productive) and outcome measure, with four levels (grammaticality judgement (GJ), selected response (SR), constrained constructed response (CCR) and translation (TR)). As a result, Study 3 was set up using a four-way 3 (Treatment) x 3 (Time) x 2 (Modality) x 4 (Outcome Measure) mixed-design template. The within-subjects variables were nested in exactly the same way as in Studies 1 and 2, with modality nested hierarchically under time, and with outcome measure nested non-hierarchically under modality. Since the four outcome measures were once again split across the two modalities, with GJ and SR nested under the receptive level of the modality variable and CCR and TR nested under the productive level of the modality variable, it was inappropriate to conduct an analysis including all independent variables. As a result, we decided once again to carry out two separate batches of analyses (see Section 6.3).
dependent variable

1

continuous

independent variables

4

categorical

written performance on a 32-item tense-related test consisting of four separate 8-item outcome measures of either a receptive or a productive nature: (1) grammaticality judgement, (2) selected response, (3) constrained constructed response and (4) translation

Between-groups variable:
1) Treatment (3 levels)
   1) Input practice
   2) Output practice
   3) No treatment

Within-groups variable:
1) Time (3 levels)
   1) Pretest
   2) Immediate posttest
   3) Delayed posttest
2) Modality (2 levels)
   1) Receptive
   2) Productive
3) Outcome Measure (4 levels)
   1) Grammaticality judgement
   2) Selected response
   3) Constrained constructed response
   4) Translation

Figure 6.1. Design box for Study 3 (3 x 3 x 2 x 4 mixed design)

6.2.2 Participants
The same population of ESL learners as the ones under investigation in Studies 1 and 2 was used for the selection of a sample of participants for Study 3. Once again, the only difference was that the participants were new and had not participated in Studies 1 or 2. For any additional information on the criteria used for the selection of participants, the reader is advised to consult Chapter 5, Section 5.2.2, since no changes were introduced for this aspect in the setup of Study 3.

Attrition. For Study 3, 101 participants initially took part in the pretest but the number of participants dropped to 94 for the immediate posttest and to 76 for the delayed posttest. In other words, at the end of the five experimental sessions we were left with the data from 76 participants. This drop corre-
sponded with an attrition rate of approximately 25%. Once again, the drop in the number of participants was the result of extraneous factors (see Chapters 4 and 5, Sections 4.2.2 and 5.2.2, Attrition). However, in comparison with the lower attrition rates for Studies 1 and 2—12.5% and approximately 10% respectively—the attrition rate for Study 3 was relatively high. An important cause for this large drop was the inclusion of a delayed posttest, which extended the duration of Study 3 by eleven weeks. In turn, this increased the possibility of higher attrition.

For the analysis and discussion of the data, the reported 25% attrition rate for Study 3 was taken into account. We tried to avoid any ceiling effect from influencing the data by screening the pretest data and excluding—as we did for Studies 1 and 2—any participants who scored more than 85% of the pretest items correctly. In practice, 85% on the pretest corresponded with a score of 28 or higher on a possible maximum score of 32. After screening the data, it was found that only one participant met the criteria for exclusion and the data from this one participant were not taken into consideration. After taking into consideration the attrition rate and the screening procedure, 75 participants remained. The data from these 75 participants were considered in the descriptive and inferential data analyses.

Pretest results also showed that there were three participants whose pretest scores (10, 12 and 13 on a possible maximum score of 32) were much lower than the mean pretest score (20). All three of these participants had been randomly assigned to the no-treatment group prior to carrying out the pretest. In practice, we could have avoided such low scores from being analysed by using a screening procedure to detect any floor effects. However, since the attrition rate for Study 3 was already high, we decided to run analyses with and without the data from these participants. Because the results of both sets of analyses led to the same conclusions, we decided to retain the data provided by the three low-scoring participants.

**Assignment to treatment groups.** The participants were assigned to treatment groups in the same way as in Studies 1 and 2. This resulted in random assignment to one of the two treatment groups and to the no-treatment group as follows: 28 (input practice), 26 (output practice) and 21 (no treatment).

**ESL grammar course.** The experimental sessions were once again conducted as part of the participants’ regular English Grammar course in the first year of their English bachelor programme. The only difference was that Study 3 took place in the academic year 2009-2010, which was two years after Study 1 and one year after Study 2. All the other course-related information was the same as the information provided for Studies 1 and 2 (see Chapters 4 and 5, Sections 4.2.2 and 5.2.2, ESL grammar course).

**ESL language proficiency.** In line with the criteria used for Studies 1 and 2, the participants’ overall level of proficiency for English was categorized as
were once again dependent on the participants’ own choices within the constraints of the bachelor programme structure in use. The context in which the foreign language combinations were made was described in detail in Chapter 4, Section 4.2.2, Academic exposure to other L2s. The descriptives for the 75 participants’ foreign language combinations found in Study 3 are presented in Table 6.1.

### Table 6.1. Participants’ language combinations (Study 3)

<table>
<thead>
<tr>
<th>English</th>
<th>other L2(s)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>Spanish</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Chinese (Mandarin)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Italian</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Portuguese</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Russian</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>French</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><strong>Subtotal</strong></td>
<td><strong>65</strong></td>
</tr>
<tr>
<td>yes</td>
<td>French and Italian</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>French and Spanish</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>French and German</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>French and Russian</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>German and Russian</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Subtotal</strong></td>
<td><strong>10</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Overall total</strong></td>
<td><strong>75</strong></td>
</tr>
</tbody>
</table>

Of the 75 participants, 65 had selected English as part of an L2 combination consisting of two L2s. As far as instruction was concerned, these 65 participants received seven hours of ESL instruction on a weekly basis. For the other L2 in their L2 combination they also received seven hours of instruction on a weekly basis. As was the case for all of the other first-year participants, these 65 participants were expected to take Dutch language courses since Dutch was their A language. The remaining ten participants had chosen English as part of
an L2 combination consisting of three L2s. These participants received either six or seven hours of English on a weekly basis, which they combined with two other L2s and Dutch (as their A language).

**Gender and age.** Of the 75 participants, 15 were male and 60 were female. The participants ranged in age from 17 to 28 years and the mean age was 20 years ($SD = 2$, mode $= 19$, median $= 19$).

**L1 backgrounds.** As was the case for the participants in Studies 1 and 2, the participants in Study 3 were asked about their L1 backgrounds and about which language(s) that they themselves regarded as their L1(s). Of the 75 participants, 73 indicated that they had one L1. Of those 73 participants, 65 participants had Dutch as their L1. The remaining eight participants’ L1s were distributed as follows, with the number of participants in parentheses: Chinese (1), Persian (1), Polish (1), Portuguese (2), Romanian (1) and Russian (2). Two participants indicated that they had two L1s, which were distributed as follows: Dutch–French and Dutch–Turkish.

**Remuneration.** The same remuneration information as that provided for Studies 1 and 2 was applied for Study 3.

### 6.2.3 Apparatus and materials

#### 6.2.3.1 Apparatus

The participants completed all of the five experimental sessions in the same computer language lab that was used for Studies 1 and 2. In addition, the infrastructure in the language lab was the same as that found in Studies 1 and 2. The only difference were the updated computers and computer programs which the participants used for the experimental sessions. Since these updates were of no significance for the actual execution of the individual experimental sessions, no additional update-related information will be provided.

For Study 3, the computer specialist who was contracted for Studies 1 and 2 was once again asked to oversee the programming of the software. Apart from some design-related changes (see Section 6.2.1), no fundamental changes were made to the programming of the experimental sessions. Technical support was once again available throughout the entire duration of the five experimental sessions and no major computer-related and/or network-related problems were reported.

#### 6.2.3.2 Materials

For a detailed description of the materials used for Study 3, the reader is advised to read Chapter 4, Section 4.2.3.2, which provides in-depth information about the materials used. Since most of this information did not change for Study 3, the description provided in this section will be limited to any changes that were made to the materials. The changes will be discussed in the following subsections: pretest, treatment theory, treatment practice, immediate posttest
and delayed posttest. This subdivision reflects the setup of the five experimental sessions in Study 3.

**Pretest.** The corpus of texts which was used for Studies 1 and 2 was also used for Study 3. For information related to the creation of the text corpus, see Chapter 4, Section 4.2.3.2, Pretest. The exact same pretest which was used for Study 2 was used for Study 3. For additional pretest information, see Chapter 5, Section 5.2.3.2, Pretest.

**Treatment theory.** No changes were made to the treatment theory. All of the participants who were assigned to one of the two treatment groups (input practice, output practice) received the exact same theoretical information as the participants in Study 2. For detailed information on the theory session, see Chapter 4, Section 4.2.3.2, Treatment theory. In addition to having two treatment groups, we decided to include a no-treatment group in Study 3. In effect, the participants in the no-treatment group received no theoretical information about the past/present perfect distinction in present-day English at all. In fact, the participants in the no-treatment group received no tense-related information whatsoever. However, they did receive theoretical information about the grammatical category of number (with respect to nouns) in present-day English since we wanted the participants to take part in all of the five experimental sessions set up for Study 3. Since the nature of the treatment theory for the no-treatment group did not directly address any of the research questions under investigation and was not linked to the target features under investigation, no further information about this topic will be provided here.

**Treatment practice.** For Study 3, the treatment practice for the two experimental groups was identical with the practice provided in Study 2. For information about the nature of the practice, see Chapter 5, Section 5.2.3.2, Treatment practice. The treatment practice provided to the participants in the no-treatment group consisted of grammatical exercises relating to the grammatical category of number (e.g., (un)countability, pluralization), which was the focus of the treatment theory for the no-treatment group. Once again, the nature of the treatment theory for the no-treatment group did not directly address any of the research questions under investigation and was not linked to the target features under investigation, so no further information about this topic will be provided here.

**Immediate posttest.** One week after completing the treatment sessions, all of the participants were asked to take part in an unannounced immediate posttest. The setup of the posttest in Study 3 was by and large identical with the setup of the pretest. For any further information, see Chapter 5, Section 5.2.3.2, Posttest. The only minor change was that we included eight short questions about the grammatical category of number in present-day English to accommodate the participants in the no-treatment group. The questions about number were of a grammaticality-judgement, selected-response and constrained-
CHAPTER 6

constructed-response nature but since they are not relevant to the research at hand, no further information will be provided about these questions here.

**Delayed posttest.** Eleven weeks after completing the immediate posttest, the participants were invited to participate in an unannounced delayed posttest. The participants were not informed that the session was in any way related to the experimental session that they had completed eleven weeks before. The delayed posttest was identical with the immediate posttest. No changes were made and the participants were simply given the same instructions as the ones that they had received for the immediate posttest.

### 6.2.4 Procedure

Study 3 was organized in the months of November, December and February of the academic year 2009-2010. It consisted of five experimental sessions, which were divided as follows: pretest (one session), treatment theory (one session), treatment practice (one session), immediate posttest (one session) and delayed posttest (one session). The first four sessions were consecutive in that they followed each other with a week between each session. These sessions all took place in November and December of 2009. The delayed posttest—which was the fifth and final session—took place in February of 2010, that is, eleven weeks after the immediate posttest. For any additional procedural information, see Chapter 5, Section 5.2.4. During the course of Study 3—as was the case with Studies 1 and 2—the participants were not informed about the setup of the experimental sessions. In other words, the participants did not know in advance what they would be doing during any of the experimental sessions. What the participants knew was that sessions would be taking place. At the beginning of every session, the participants received the relevant information for that specific session.

### 6.2.5 Coding and scoring

No changes were made to the coding and scoring procedure. In other words, the procedures explained for Studies 1 and 2 in Chapters 4 and 5, Sections 4.2.5 and 5.2.5, were replicated and used for Study 3.

### 6.2.6 Statistical analyses

With regard to the statistical analyses, the same logic that was applied to the design in Studies 1 and 2 was applied to the design in Study 3. Instead of running a four-way 3 (Treatment) x 3 (Time) x 2 (Modality) x 4 (Outcome Measure) RM ANOVA, we decided to run two separate batches of analyses (see Section 6.3). For any other details concerning the statistical analyses, the reader is advised to read Chapter 4, Section 4.2.6.
6.2.7 Validity and reliability
Since we used the same corpus of texts for the selection of pretest and posttest items, the information provided in Chapter 4, Section 4.2.7, on the validity of the test items also applies to the texts which were used for Study 3. Reliability was once again measured using Cronbach’s alpha coefficient (α). The Cronbach’s alpha coefficient for the pretest was .54. The Cronbach’s alpha coefficients for the immediate posttest and delayed posttest were .80 and .79 respectively.

6.3 Results
The results reported in this section are all based on the quantitative data, which were collected throughout the pretest and posttest sessions.

Once again, analyses for both main effects and interaction effects were carried out at all of the three levels of the experimental design as indicated in Figure 4.9 in Chapter 4. On the whole, only those results relevant in light of the research questions that were formulated in Section 6.1 have been highlighted and discussed below.

The first batch of analyses consisted of two three-way ANOVAs, one 3 (Treatment) x 2 (Time) x 2 (Modality) RM ANOVA and one 3 (Treatment) x 3 (Time) x 2 (Modality) RM ANOVA. For both RM ANOVAs, treatment was the between-subjects variable and time and modality the within-subjects variables. For the first RM ANOVA we did not consider the delayed posttest results since we first wanted to ascertain whether a treatment effect was visible on the immediate posttest. The second RM ANOVA, however, did include the delayed posttest results and was run to analyse the durability of any possible treatment effects that were found. The descriptives are shown numerically in Table 6.2 and graphically in Figures 6.2, 6.3, 6.4 and 6.5.
Table 6.2: Performance on receptive and productive pretest and posttest items by treatment group
(maximum scores of 16 on receptive and productive items and 32 on the totality of pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Modality</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI_{95%}</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>10.21</td>
<td>2.15</td>
<td>9.38</td>
<td>11.05</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>10.00</td>
<td>2.19</td>
<td>9.12</td>
<td>10.88</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>8.00</td>
<td>2.19</td>
<td>7.00</td>
<td>9.00</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>9.52</td>
<td>2.35</td>
<td>8.98</td>
<td>10.06</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>10.92</td>
<td>1.57</td>
<td>10.29</td>
<td>11.56</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>9.81</td>
<td>2.11</td>
<td>8.85</td>
<td>10.77</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>10.31</td>
<td>1.93</td>
<td>9.86</td>
<td>10.75</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>20.32</td>
<td>3.52</td>
<td>18.96</td>
<td>21.69</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>20.92</td>
<td>3.15</td>
<td>19.65</td>
<td>22.19</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>17.81</td>
<td>3.76</td>
<td>16.10</td>
<td>19.52</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>19.83</td>
<td>3.66</td>
<td>18.99</td>
<td>20.67</td>
<td>10</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; LL = lower limit; UL = upper limit.
Table 6.2. (Contd.), Performance on receptive and productive pretest and posttest items by treatment group
(maximum scores of 16 on receptive and productive items and 32 on the totality of pretest or posttest items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Modality</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI</th>
<th>Min</th>
<th>Max</th>
<th>L.L.</th>
<th>U.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Receptive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>12.79</td>
<td>1.93</td>
<td>12.04</td>
<td>13.53</td>
<td>9</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>13.15</td>
<td>1.91</td>
<td>12.38</td>
<td>13.93</td>
<td>8</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>9.19</td>
<td>2.66</td>
<td>7.98</td>
<td>10.40</td>
<td>4</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>11.91</td>
<td>2.72</td>
<td>11.28</td>
<td>12.53</td>
<td>4</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>12.57</td>
<td>2.35</td>
<td>11.66</td>
<td>13.48</td>
<td>7</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>11.92</td>
<td>2.12</td>
<td>11.07</td>
<td>12.78</td>
<td>9</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>8.71</td>
<td>2.81</td>
<td>7.43</td>
<td>9.99</td>
<td>3</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>11.27</td>
<td>2.88</td>
<td>10.60</td>
<td>11.93</td>
<td>3</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Receptive &amp; productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>25.36</td>
<td>3.52</td>
<td>23.99</td>
<td>26.72</td>
<td>19</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>25.08</td>
<td>3.45</td>
<td>23.68</td>
<td>26.47</td>
<td>17</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>17.90</td>
<td>4.28</td>
<td>15.96</td>
<td>19.85</td>
<td>10</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>23.17</td>
<td>4.94</td>
<td>22.04</td>
<td>24.31</td>
<td>10</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; L.L. = lower limit; U.L. = upper limit.

The first RM ANOVA yielded a significant main effect of time, $F(1, 72) = 75.370$, $p = .000$, $\eta^2_p = .511$. In other words, performance on all 32 items showed a statistically significant, large increase from 19.83 in the pretest to 23.17 in the immediate posttest. In addition, we also found a significant main effect of treatment, $F(2, 72) = 18.288$, $p = .000$, $\eta^2_p = .337$. No main effect of
modality was obtained. In addition to finding significant main effects, we also found interactions. Of special interest to us was the Time x Treatment interaction, which was significant, $F(2, 72) = 16.975, p = .000, \eta^2_p = .320$ when comparing the pretest and immediate posttest scores (see Figure 6.4). Tukey HSD post-hoc tests revealed that the input-practice group scored significantly higher on the immediate posttest than the no-treatment group (mean difference = 2.49, CI95% = 1.37–3.62, $p = .000$) and that the output-practice group also scored significantly higher on the same test than the no-treatment group (mean differences = 2.57, CI95% = 1.43–3.71, $p = .000$). There were no significant differences between the input-practice and output-practice groups (mean differences = -0.08, CI95% = -1.14–0.98, $p = .982$). Another interaction which was found to be statistically significant was the Time x Modality interaction, $F(1, 72) = 15.581, p = .003, \eta^2_p = .178$ (see Figure 6.5). However, since the mean scores differ only marginally across time and modality, we attribute no importance to the Time x Modality interaction in light of the research questions under investigation.

When taking into consideration the delayed posttest data, we found similar results. The RM ANOVA also yielded a significant main effect of time, $F(2, 144) = 26.357, p = .000, \eta^2_p = .268$. In addition, we also found a main effect of treatment, $F(2, 72) = 15.193, p = .000, \eta^2_p = .297$. Once again there were interactions which showed statistical significance. The Time x Treatment interaction was significant, $F(4, 144) = 22.648, p = .000, \eta^2_p = .149$ (See Figure 6.4). Tukey HSD post-hoc tests once again revealed that the input-practice group scored significantly higher on the tests as a whole than the no-treatment group (mean difference = 2.42, CI95% = 1.27–3.58, $p = .000$) and that the output-practice group also scored significantly higher on the tests as a whole than the no-treatment control group (mean differences = 2.31, CI95% = 1.13–3.48, $p = .000$). There were no significant differences between the input-practice and output-practice groups (mean differences = 0.12, CI95% = -0.98–1.21, $p = .965$). The Time x Modality interaction was once again statistically significant, $F(2, 144) = 9.074, p = .000, \eta^2_p = .112$ (see Figure 6.5) but since the mean scores differ only marginally across time and modality, we attribute no importance to the Time x Modality interaction in light of the research questions under investigation.
Figure 6.2. Performance on receptive pretest, immediate posttest and delayed posttest by treatment group (maximum scores of 16)
Figure 6.3. Performance on productive pretest, immediate posttest and delayed posttest by treatment group (maximum scores of 16)
Figure 6.4. Interaction between time of testing and treatment (maximum scores of 32)

Figure 6.5. Interaction between time of testing and modality (maximum scores of 16)
Subsequently, we conducted a second batch of analyses, consisting of two three-way ANOVAs: one $3$ (Treatment) x $2$ (Time) x $4$ (Outcome Measure) RM ANOVA and one $3$ (Treatment) x $3$ (Time) x $4$ (Outcome Measure) RM ANOVA. For both RM ANOVAs Treatment was still the between-subjects variable and time and outcome measure the within-subjects variables. For the first RM ANOVA, we did not consider the delayed posttest results since we first wanted to ascertain whether a treatment of effect was visible on the immediate posttest. The second RM ANOVA, however, did include the delayed posttest results and was run to analyse the durability of any possible treatment effects that were found. The descriptives are shown numerically in Table 6.3 and graphically in Figures 6.6 and 6.7.

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI 95%</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td>LL</td>
<td>UL</td>
<td></td>
</tr>
<tr>
<td>PRETEST</td>
<td>GJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>28</td>
<td>4.57</td>
<td>1.48</td>
<td>4.00</td>
<td>5.14</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Output practice</td>
<td>26</td>
<td>4.27</td>
<td>1.43</td>
<td>3.69</td>
<td>4.85</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>No treatment</td>
<td>21</td>
<td>3.29</td>
<td>1.65</td>
<td>2.54</td>
<td>4.04</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>75</td>
<td>4.11</td>
<td>1.58</td>
<td>3.74</td>
<td>4.47</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>SR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>28</td>
<td>5.64</td>
<td>1.34</td>
<td>5.12</td>
<td>6.16</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Output practice</td>
<td>26</td>
<td>5.73</td>
<td>1.51</td>
<td>5.12</td>
<td>6.34</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>No treatment</td>
<td>21</td>
<td>4.71</td>
<td>1.23</td>
<td>4.15</td>
<td>5.27</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>75</td>
<td>5.41</td>
<td>1.43</td>
<td>5.09</td>
<td>5.74</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>CCR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>28</td>
<td>5.79</td>
<td>1.20</td>
<td>5.32</td>
<td>6.25</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Output practice</td>
<td>26</td>
<td>5.85</td>
<td>1.12</td>
<td>5.39</td>
<td>6.30</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>No treatment</td>
<td>21</td>
<td>5.48</td>
<td>1.29</td>
<td>4.89</td>
<td>6.06</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Mean</td>
<td>75</td>
<td>5.72</td>
<td>1.19</td>
<td>5.45</td>
<td>5.99</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>TR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input practice</td>
<td>28</td>
<td>4.32</td>
<td>1.49</td>
<td>3.74</td>
<td>4.90</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Output practice</td>
<td>26</td>
<td>5.08</td>
<td>1.32</td>
<td>4.54</td>
<td>5.61</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>No treatment</td>
<td>21</td>
<td>4.33</td>
<td>1.28</td>
<td>3.75</td>
<td>4.92</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Mean</td>
<td>75</td>
<td>4.39</td>
<td>1.41</td>
<td>4.26</td>
<td>4.91</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.

Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Table 6.3. (Contd.). Performance on pretest and posttest items by outcome measure and by treatment group
(maximum scores of 8 on every outcome measure and 32 on all test items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>Treatment</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI&lt;sub&gt;95%&lt;/sub&gt;</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMMEDIATE POSTTEST</td>
<td>GJ</td>
<td>Input practice</td>
<td>28</td>
<td>6.32</td>
<td>1.02</td>
<td>5.93 - 6.72</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>6.35</td>
<td>1.33</td>
<td>5.81 - 6.88</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>4.52</td>
<td>1.47</td>
<td>3.85 - 5.19</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>75</td>
<td>5.83</td>
<td>1.49</td>
<td>5.48 - 6.17</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>SR</td>
<td>Input practice</td>
<td>28</td>
<td>6.46</td>
<td>1.26</td>
<td>5.98 - 6.95</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>6.81</td>
<td>.94</td>
<td>6.43 - 7.19</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>4.67</td>
<td>1.80</td>
<td>3.85 - 5.49</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>6.08</td>
<td>1.60</td>
<td>5.71 - 6.45</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>CCR</td>
<td>Input practice</td>
<td>28</td>
<td>6.61</td>
<td>1.03</td>
<td>6.21 - 7.01</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>6.46</td>
<td>6.02</td>
<td>6.91 - 1.10</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>4.90</td>
<td>1.81</td>
<td>4.08 - 5.73</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>6.08</td>
<td>1.50</td>
<td>5.74 - 6.42</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>Input practice</td>
<td>28</td>
<td>5.96</td>
<td>1.88</td>
<td>5.24 - 6.69</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>5.46</td>
<td>1.70</td>
<td>4.77 - 6.15</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>3.81</td>
<td>1.47</td>
<td>3.14 - 4.48</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>75</td>
<td>5.19</td>
<td>1.91</td>
<td>4.75 - 5.63</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.
Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).
Table 6.3 (Contd.). Performance on pretest and posttest items by outcome measure and by treatment group
(maximum scores of 8 on every measure and 32 on all test items)

<table>
<thead>
<tr>
<th>Time</th>
<th>Outcome Measure</th>
<th>Treatment</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>CI 95%</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>GJ</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>6.04</td>
<td>1.43</td>
<td>5.48</td>
<td>6.59</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>5.62</td>
<td>1.47</td>
<td>5.02</td>
<td>6.21</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>4.76</td>
<td>1.30</td>
<td>4.14</td>
<td>5.35</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>75</td>
<td>5.53</td>
<td>1.48</td>
<td>5.19</td>
<td>5.87</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>SR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>5.89</td>
<td>1.52</td>
<td>5.30</td>
<td>6.48</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>6.15</td>
<td>2.05</td>
<td>5.32</td>
<td>6.98</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>4.76</td>
<td>1.04</td>
<td>4.29</td>
<td>5.24</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>75</td>
<td>5.67</td>
<td>1.70</td>
<td>5.27</td>
<td>6.06</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>CCR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>6.50</td>
<td>1.14</td>
<td>6.06</td>
<td>6.94</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>6.04</td>
<td>1.54</td>
<td>5.42</td>
<td>6.66</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>5.95</td>
<td>1.40</td>
<td>5.32</td>
<td>6.59</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>75</td>
<td>6.19</td>
<td>1.36</td>
<td>5.87</td>
<td>6.50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Input practice</td>
<td>28</td>
<td>5.39</td>
<td>1.81</td>
<td>4.69</td>
<td>6.10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Output practice</td>
<td>26</td>
<td>5.00</td>
<td>2.10</td>
<td>4.15</td>
<td>5.85</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No treatment</td>
<td>21</td>
<td>3.76</td>
<td>2.05</td>
<td>2.83</td>
<td>4.69</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>75</td>
<td>4.80</td>
<td>2.07</td>
<td>4.32</td>
<td>5.28</td>
<td>1</td>
</tr>
</tbody>
</table>

Note 1. CI = confidence interval; LL = lower limit; UL = upper limit.
Note 2. The within-subjects variable outcome measure was defined as having four levels: (1) grammaticality judgement (GJ), (2) selected response (SR), (3) constrained constructed response (CCR) and (4) translation (TR).

The first batch of RM ANOVAs yielded a significant main effect of time, $F(1, 216) = 75.370, p = .000, \eta^2_p = .511$, which was the same effect that we had already discovered in the first of the previous batch of RM ANOVAs. Apart from the reported significant main effect of time, a significant main effect of outcome measure was also found, $F(3, 216) = 24.570, p = .000, \eta^2_p = .254$. In addition to finding significant main effects, we once again found interactions. The Time x Treatment interaction reported above was once again significant, $F(2, 216) = 16.975, p = .000, \eta^2_p = .320$ (see Figure 6.4). Tukey HSD post-hoc tests revealed, as they did before, that the input-practice treatment groups scored significantly higher on the immediate posttest than the no-treatment group (mean difference = 1.25, CI<sub>95%</sub> = 0.68–1.81, $p = .000$) and that the output-practice treatment group also scored significantly higher on the same test than the no-treatment group (mean differences = 1.29, CI<sub>95%</sub> = 0.71–1.86, $p =$
.000). There were no significant differences between the input-practice and output-practice treatment groups (mean differences = -0.04, CI95% = -0.57–0.49, p = .982). Another interaction which was found to be significant was the Time x Outcome Measure interaction, F(3, 216) = 10.558, p = .000, η²p = .128. However, since the mean scores differ only marginally across time and modality, we attribute no importance to the Time x Modality interaction in light of the research questions under investigation.

When taking into consideration the delayed posttest data, we found similar results. The second batch of RM ANOVAs also yielded a significant main effect of time, F(2, 432) = 26.357, p = .000, η²p = .268. Moreover, we also found a main effect of treatment, F(2, 72) = 15.193, p = .000, η²p = .297. Once again there were interactions which showed statistical significance. The Time x Treatment interaction was significant, F(4, 432) = 6.288, p = .000, η²p = .149. Tukey HSD post-hoc tests revealed an identical picture as the one revealed above. The input-practice treatment groups scored significantly higher on the tests as a whole than the no-treatment group (mean difference = 1.21, CI95% = 0.63–1.79, p = .000). The output-practice treatment group also scored significantly higher on the tests as a whole than the no-treatment group (mean differences = 1.15, CI95% = 0.57–1.74, p = .000). There were no significant differences between the input-practice and output-practice treatment groups (mean differences = 0.06, CI95% = -0.49–0.60, p = .965). The Time x Outcome Measure interaction was once again significant, F(6, 432) = 8.056, p = .000, η²p = .101. Both the descriptive and inferential statistics show that the four outcome measures were not equally difficult but since the mean scores differ only marginally across treatment, time and outcome measure, we attribute no importance to the main of effect of outcome measure and the Time x Outcome Measure interaction in light of the research questions which were investigated.
Figure 6.6. Pretest, immediate posttest and delayed posttest performance on all four outcome measures by treatment group (maximum scores of 8 with GJ0 = grammaticality judgement pretest, GJ1 = grammaticality judgement immediate posttest, GJ2 = grammaticality judgement delayed posttest, etc.)
Figure 6.7. Performance by time of testing and outcome measure (maximum scores of 8)
In light of the results of the analyses presented above, we did not deem it worthwhile to pursue any subsequent analyses.

6.4 Conclusions and discussion
With respect to the four research questions formulated at the beginning of this chapter, the following general conclusions can be drawn. Analyses of the data in Study 3 revealed a large increase from pretest to immediate posttest performance. This increase was a 3.34-point increase when comparing overall pretest and immediate-posttest scores (20 and 23 respectively on a total score of 32). When breaking down this increase according to treatment, it was clear to see that there was a treatment effect. Both treatment groups showed significant increases from pretest to immediate-posttest performance (20 to 25 for the input-practice group, 21 to 25 for the output-practice group). The no-treatment group showed no significant increase from pretest to immediate-posttest performance (17.81–17.90). When comparing the performances on the delayed posttest, an effect of treatment was still visible although the overall delayed-posttest scores had decreased for both the input-practice group and output-practice group (24 and 23 respectively). Apart from the clear effects of time and treatment, there were no main effects of mode and outcome measure, nor were there any interactions which indicated any differential effects of treatment.

In the next chapter (Chapter 7), a more detailed account of the results will be provided for all three of the studies carried out and reported on in Chapters 4, 5 and 6. In addition to summarizing the findings, interpretations will be provided taking into account theoretical findings highlighted in Chapters 1, 2 and 3. Chapter 7 will also feature a discussion of the strengths and limitations of the three studies that have been carried out and a discussion of the important implications for tense-related language pedagogy. By way of conclusion, suggestions will be provided for further tense-related research into the effects of instruction.
CHAPTER 7
GENERAL DISCUSSION

Practice does not make perfect. Only perfect practice makes perfect.
(Vincent T. Lombardi)

7.1 Introduction
Having provided, in Chapters 1, 2 and 3, a theoretical framework for the experimental research reported on in Chapters 4, 5 and 6, I would like to end this doctoral dissertation with a general discussion of the research in its entirety. This chapter should be regarded as a bridge between the three theoretical chapters and the three experimental chapters. The first section of this chapter (Section 7.2) will focus on the experimental research and will provide a summary of the findings across all three of the studies that were carried out. In the second section (Section 7.3), the findings from the studies will be discussed in light of the theory that was presented in the first three chapters of this dissertation. The third section (Section 7.4) will underscore some of the major strengths of the experimental research that was carried out. However, as is the case with any form of experimental research, the studies reported on in Chapters 4, 5 and 6 have limitations. These limitations will also be discussed in the third section. The fourth section (Section 7.5) will focus on the implications for language pedagogy since the effect of form-focused instruction (FFI) was the general foundation on which the research was carried out. In the fifth section of this chapter (Section 7.6), suggestions for further research will be discussed. The sixth and final section (Section 7.7) will provide some final observations, which will serve as an overall conclusion to this dissertation.

7.2 Summary of the findings
As a result of design-related and methodological changes that were made to the studies as the research progressed from one study to the next, the findings across all three of the studies do not provide completely identical pictures. However, regardless of the experimental ‘tweaks’ that were carried out, there are some clear similarities and differences, which will be highlighted and summarized below. Before providing a summary, let us, by way of a reminder, have another look at the three research questions that were formulated in Chapters 4, 5 and 6:
Research questions

1. Does explicit form-focused instruction (FFI) have an effect on (Dutch-speaking) ESL learners’ performance with regard to complex temporal form–meaning–use mappings (i.e., the past/present perfect distinction) in L2 English?
2. If explicit FFI does have an effect, can any overall differential effects be ascertained with respect to the specific type of instructional treatment (e.g., input practice, output practice)?
3. If differences between treatments can be ascertained, are the differential effects the same across all of the outcome measures? Or are there differences between receptive and productive outcome measures?

It should be reiterated that even though all three research questions were formulated in all three of the studies, research question 1 could not actually be addressed in Studies 1 and 2 since no control groups were integrated into the research design of those studies. In addition to the three research questions above, there was a fourth research question, which was formulated only in Study 3. This fourth research question sought to address the durability of any treatment effects and was formulated as follows:

4. Are any treatment effects of a durable nature?

Table 7.1 provides an overview of which research questions were addressed in which of the studies. Although increases in test performances (from pretest to posttest) were observed in both Studies 1 and 2, we were unable to state—without any available data from control groups for statistical analyses—that these increases resulted from the treatments that had been provided. Consequently, we refrained from addressing research question 1 altogether in Studies 1 and 2. We assumed that the treatment would have an effect based on existing findings about the effects of instruction in SLA research.

Table 7.1. Research questions addressed across Studies 1, 2 and 3
(Yes = research question addressed, no = research question not addressed)

<table>
<thead>
<tr>
<th>Study</th>
<th>Research question 1</th>
<th>Research question 2</th>
<th>Research question 3</th>
<th>Research question 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Comparisons of the overall pretest and posttest performances for Studies 1 and 2, and the overall pretest and immediate-posttest performances for Study 3,
showed that in all three of the studies significant increases in test performances from pretest to (immediate) posttest were observed under treatment conditions. In Study 1, the increase was relatively small (approximately 1.5 on a total test score of 24) (see Table 4.3 and Figures 4.10, 4.11 and 4.12 in Chapter 4), whereas in Studies 2 and 3 the increases were relatively large (approximately 4.5 and 3.5 respectively on a total test score of 32) (see Table 5.2 and Figures 5.3, 5.4 and 5.5 in Chapter 5, Table 6.2 and Figures 6.2 and 6.3 in Chapter 6). As stated above, it was impossible to state that the observed increases in Studies 1 and 2 were a result of any treatment effects since no control groups had been integrated into the research design. However, this was not the case for Study 3. For Study 3, a no-treatment (control) group was integrated into the experimental setup, with the participants in the no-treatment group not receiving any tense-related instruction at all during the experimental sessions. By contrast, the two treatment groups in Study 3 (input practice, output practice) both received tense-related instruction. When comparing the overall pretest and immediate-posttest performances for the three groups of participants in Study 3 (input practice, output practice, no treatment), we found significant treatment effects. Both the input-practice group and the output-practice group showed significant increases in pretest and immediate-posttest performances (approximately 5 and 4 respectively on a total test score of 32) (see Table 6.2). The no-treatment group, however, did not show any significant increase (see Table 6.2). A comparison of the delayed-posttest performances for Study 3 indicated that an effect of treatment was still observable even though initial test-score increases on the immediate-posttest had decreased somewhat when we analysed the overall delayed-posttest scores for both the input-practice and the output-practice groups (4 and 3 respectively, on a total test score of 32) (See Table 6.2). In other words, the treatment effect found on the immediate-posttest performances was still visible and statistically significant on the delayed-posttest performances but the durability of the treatment effect had been affected in a slightly negative way. All of these findings help in answering the first research question for Study 3 as follows: The explicit FFI provided to the participants in the input-practice and output-practice groups did indeed have a positive effect on the participating ESL learners’ test performances with regard to the complex temporal FMU mapping under investigation.

With regard to the second research question, the following can be said: As far as the differential effects were concerned with respect to the specific type of instructional treatment (input practice, output practice), no significant overall differences were found in any of the three studies between the various treatment groups.

As far as the third research question is concerned, we did not find differential effects on receptive and productive outcome measures. Analyses of any interaction between treatment and modality did not provide any significant results. In other words, no differential effects were found between treatment—
input practice and output practice—and receptive and productive outcome measures.

Thus, the results of the three studies may be summarized as follows: Explicit FFI was found to have a durable positive effect on participants’ posttest performances but no differential effects were found between input-practice and output-practice groups, and between treatments and receptive and productive outcome measures.

The next section will turn to a discussion of the findings within the broader theoretical context provided in Chapters 1, 2 and 3.

7.3 Discussion of the findings

As far as research questions 1 and 4 are concerned, the findings are in line with general findings on the effects of explicit, form-focused instruction (FFI) in contemporary SLA research (see, for example, Norris & Ortega, 2000, and the literature presented and discussed in Chapter 3, Section 3.3). The assumption about the positive, instructional effects explains the directional research hypothesis for research question 1, which stated that explicit FFI instruction would have a positive effect on the acquisition of the L2 target feature under investigation. Although we observed significant differences in the participants’ test performances when comparing pretest and posttest performances and when comparing treatment groups with no-treatment groups, two questions remain with respect to the effect of instruction: (1) Did the FFI instruction provided actually bring about changes in the ESL learners’ competence or in their performance? and (2) Did the data-collection instruments in fact collect competence-related data or performance-related data? Direct insights into L2 learners’ competence are notoriously difficult to gain. Instead, most researchers are forced to rely on L2 learners’ performances in the hope that they will be able to gain any insights to what they believe to be competence (see, for example, Shohamy, 1996, for a discussion of the competence/performance distinction for language testing). Since we were not directly concerned with the competence/performance distinction for the studies carried out for this doctoral dissertation, this is not an issue that we explicitly addressed in the research. Consequently, this is not an issue which will be discussed any further there.

As far as skill specificity is concerned, the picture that this dissertation provides is somewhat more intricate, as explained in Chapter 3, Section 3.3.2. Both input processing (IP)/processing instruction (PI) and skill learning theory have different views on skill specificity and that is why non-directional research hypotheses were consciously formulated when drawing up research questions 2 and 3. Non-directional hypotheses allowed us to test the predictions of both approaches. Following IP/PI predictions, the acquisition of L2 target features was not expected to be skill-specific. In the context of the three studies that were carried out, IP/PI proponents would have predicted what Lee and Benati
(2007) state in their comparative discussions of studies into the effects of PI, of traditional instruction (TI) and of meaning-based output instruction (MOI):

The findings are robust that PI > TI on interpretation tasks. Learners who receive PI score significantly higher on interpretation tasks than learners who receive TI, no matter the target language nor the processing problem addressed. The findings are also robust that PI = TI on form production tasks. Learners who receive PI, during which they never produce a target form, score the same on form production tasks no matter the target language nor the processing problem addressed. (p. 40)

The findings of these studies on learners’ performance on form production tasks is perfectly consistent. Learners who receive PI improve to the same extent as those who receive MOI on form production in that there is no significant difference between scores for the PI and MOI groups. The findings of this line of investigation on the interpretation task are not consistent. Those learners who receive PI sometimes score significantly higher on interpretation tasks than those who receive MOI ... and sometimes make equal gains ... (p. 43)

In other words, if the input (practice)-based instruction provided in the studies qualifies as processing instruction, IP/PI proponents predict that the three experimental studies should indicate that the input-practice group will perform better than the output-practice group on receptive tasks and (at least) the same on productive tasks.

Of special interest in this respect is the second quote above, which aligns with the output-based treatments provided in the three studies. In the quote, it is clearly stated that some studies have indicated that there are times when L2 learners who receive PI and L2 learners who receive MOI make equal gains on interpretation tasks (i.e., comprehension tasks or receptive tasks). However, on the back of this statement, Lee and Benati (2007) are quick to add that the findings that equal gains are sometimes made on interpretation tasks—when comparing PI with MOI—are isolated and exceptional findings. In addition, they add that the studies which report such findings (e.g., Farley, 2001b; Morgan-Short & Bowden, 2006) may be further isolated with new research, which confirms the PI > MOI result with respect to interpretation tasks (Benati & Lee, 2007). Farley (2001b) initially provided an explanation for the PI = MOI results on interpretation tasks in his research by stating that the MOI treatment group may, for example, have been given incidental input (i.e., input made available incidentally as a result of, for example, communication between participants) because of the mode of delivery of the instruction, which was in-classroom delivery. Lee and Benati (2007) report on a study which investigates the effects of PI and MOI on the acquisition of the Italian subjunctive of doubt and opinion and the French subjunctive of doubt. As a result of using classroom deliv-

---

97 Meaning-based output instruction (MOI) consists of structured output activities, which may be regarded as the output-based counterpart of structured input activities (SIAs) (see, for example, Lee & VanPatten, 2003).
ery and computer delivery, a comparison could be made to ascertain any possible facilitative role for incidental input. The results for the computer-delivery instruction showed a PI > MOI advantage on interpretation tasks. In other words, Farley’s (2001b) statement about the provision of incidental input and his finding that PI = MOI on interpretation tasks was isolated even further.

However, with respect to the studies carried out for this dissertation, we are left with a PI = MOI result for comprehension/input-related tasks even though instruction was provided via computers in all three studies. Unlike the treatments in Farley’s (2001b) study, the treatments in Studies 1, 2 and 3 could not possibly have allowed any incidental input to result from in-classroom interaction between learners as referred to in Farley’s study. Yet, the PI participants (input practice in Studies 1, 2 and 3) did not outperform the MOI participants (output practice in Studies 1, 2 and 3) on receptive tasks. As a matter of fact, they performed equally well. It appears then that no ready-made explanation for the findings in Studies 1, 2 and 3 can be provided if we adopt an IP/PI approach.

What if we approached the findings using a skill acquisition theory approach? With skill acquisition theory, we see a different picture altogether. Following the tenets from skill learning theory, progression through the developmental stages (i.e., from declarative to procedural and from procedural to automatic) is skill-specific. Consequently, predictions about the differential effects of instruction would spell out increases for the input-practice group on receptive tasks and increases for the output-practice group on productive tasks. These increases were indeed confirmed but in addition to skill-specific increases the practice-based treatment groups in Studies 1, 2 and 3 also showed increases on the other modality. In other words, the input-practice group also increased on the productive tasks and the output-practice group on the receptive tasks without significant interaction effects between treatment and modality. This could point towards the development of skills which is either not skill-specific or only partially skill-specific depending on how one views the findings in Studies 1, 2 or 3.

It appears then that we are in a quandary with respect to the predictions made by both IP/PI and skill learning theory since both sets of predictions were not completely confirmed in the three studies. A simple explanation may be that the combination of the target features under investigation in Studies 1, 2 and the practice provided to the participants did not result in any differential effects of treatment. However, it may also be that the situation in the studies is more complex and requires further investigation with new, promising lines of research to shed light on the phenomenon. I would like to turn to the nature of skills, of skill learning and of skill specificity as a promising avenue of investigation for the findings in Studies 1, 2 and 3.

The concept of skill specificity predicts that receptive practice promotes the learning of receptive (language) skills and productive practice the learning of
productive (language) skills. What if the concept of skill specificity—as far as language skills are concerned—is not as dichotomous a concept as has been represented in the SLA literature, with, for example, PI/IP claiming no skill specificity and skill acquisition theory claiming the opposite? What if we entertain the possibility that there are positions in between ‘skill-specific’ and ‘not skill-specific’ and approached the concept of skill specificity as a continuum rather than a yes-no dichotomy? Skill specificity may simply not operate in a yes-no fashion in skill-learning contexts in general and in SLA-related skill-learning contexts in particular. As such, skill learning is recognized as a complex phenomenon and theories of skill learning reflect the complex nature of skill learning (e.g., Anderson, 1983, 1993, 2000). Why then would we expect a question about the skill specificity of developing language skills to be a question that could be answered in a straightforward manner, that is, in either the affirmative or in the negative and without any further nuances? I believe that much more about language skill learning could be discovered if distinctions such as receptive–productive were complemented with additional skill-related distinctions. One such distinction, which has already been made in the literature on skilled movements, is the distinction between closed skills and open skills (e.g., Allard & Starkes, 1991; Elliott & Lyons, 1998; Schmidt & Lee, 2005). Segalowitz (2010) discusses this distinction and describes it as follows:

Closed skills are those taking place in environments that are relatively stable and where the goal of performance is to recreate as accurately as possible a physical or cognitive act that meets a particular standard or ideal form (Allard & Starkes, 1991, p. 127). Examples include motor movement activities such as weightlifting, diving, gymnastics, and cognitive activities such as mental arithmetic. In contrast, open skills take place in environments that are relatively unpredictable and where the goal of performance is to bring about some effect upon the environment (place the puck in the hockey net by overcoming the opponents’ defensive moves; capture the opponent’s king in chess), none of which involve repeating a particular motor movement or mental calculation according to some predefined standard. (pp. 66–67)

Segalowitz (2010) recognizes that many skills cannot be neatly categorized as closed or open since they may involve features of both closed and open environments.

Although Segalowitz’s description of the open/closed distinction above may appear somewhat abstract with respect to SLA, he provides the following description of how the distinction could be applied to L2 performance:

For example, L2 learners have the option of approaching the environment as essentially closed, requiring an emphasis on accuracy and precision, and involving a great deal of self-focus. In the extreme, it might be more accurate to talk about such L2 learners as reciting L2 utterances, rather than speaking spontaneously when using the language. It may be, for example, that certain types of classroom instruction encourage this way of relating to the language environment. In contrast, L2 learners can also view the environment as essentially open, requiring a focus on what communicative and social goals
are to be accomplished with other speakers. Here the aim would be to fulfill communicative intentions and achieve goals that are not focused primarily on linguistic precision. In this case it might be more accurate to talk about L2 users navigating their communicative environment, and using language as a tool for doing so. The closed and open skill stances will necessarily place different processing demands on the L2 user. (p. 67, quoted with original highlighting)

Although the distinction between closed and open skills has not actually featured and been integrated into any SLA research paradigm (yet), it could bear explanatory power in that it has obvious implications for processing demands on the L2 learner and for language learning (Segalowitz, 2010). By extension, it may also be applied to language teaching. With respect to skill specificity, for example, it may be that the crude dichotomous distinction between skill-specific and not skill-specific is an inaccurate reflection of the skill learning process if we take into consideration the nature of the product being learnt, that is, a specific language skill or a set of language skills. Research into language skills in predominantly closed environments may reveal that as a result of increased stability in closed environments, less taxing processing demands are placed on L2 learners than when they are faced with skills in predominantly open environments. In turn, predominantly closed environments may interact with lower degrees of skill specificity and predominantly open environments with higher degrees of skill specificity if skill specificity is more accurately represented as a continuum. SLA research into this intricate area is currently non-existent but may be able to provide explanatory power to specific lines of investigation in (language) skill learning theory in SLA.

How can the closed/open distinction be applied to the practice-based instruction and the tests provided in Studies 1, 2 and 3? Both the instruction and the tests are not representative of completely closed environments. However, if they were placed on the closed-open continuum of skills, they would have to be located towards the closed end of the continuum. The linguistic contexts (i.e., linguistic environments) provided to the participants in all three studies were relatively closed in that the contexts and the task types did not require a high degree of ‘navigation’ through communicative environments. However, the four outcomes measures, which were used for the practice sessions and the testing sessions, were not all closed or open to the same degree. If we compare, for example, the selected response (SR), the constrained constructed response (CCR) and the translation (TR) task types (see Chapter 4, Figures 4.3, 4.4 and 4.5) with respect to the closed-open distinction, it can be said that all three task types were relatively closed in that the participants were given the linguistic environments in which the task types were presented. In other words, the participants were helped with respect to the creation of limited linguistic environments. At the same time, however, the task types were different in fundamental ways. The SR task type provided the participants with possible answers—three in total—from which they were asked to select the finite verb form which they
believed to be the grammatically correct verb form for the non-finite verb which appeared in bold in the linguistic context. In contrast, the CCR task type did not provide the participants with any possible answers. For the CCR task types, the participants were asked to provide a finite verb form, which—according to them—was a grammatical verb form of the non-finite verb which appeared in bold in the linguistic context. From a closed/open, skill-related point of view, the CCR task type may be described as relatively more open than the SR task type with respect to filling in or selecting the correct verb form. As far as the TR task type is concerned, drawing up any concrete comparisons with the materials available in the studies is complex. The linguistic environments in which the TR task types were presented were essentially the same as the linguistic environments found in the other task types, except for the fact that the TR task types contained one sentence in Dutch. However, important information in the Dutch sentence (e.g., temporal adverbials) was just as present in Dutch as in English. Consequently, one cannot truly say that the linguistic environments were fundamentally different. The actual task type itself, however, was. This was also observed in the qualitative data gathered for Study 1 (see Chapter 4, Section 4.6). The TR task type was considered by many participants as difficult in that they felt that they were required to do more than in the SR and CCR task types. Not only did they have to find the appropriate words, expressions, idioms but they were also asked to pay attention to, for example, grammar, syntax and style. In effect, many participants felt that their attentional resources were taxed more in the TR task types than in any other of the task types. For all the inherent complexity in the past/present perfect distinction (see Chapter 2), there are also other factors (e.g., linguistic context, task complexity) which may contribute to increased complexity. At a sentential level, the past/present perfect may be viewed as less challenging for L2 learners than at a suprasentential level since L2 learners are generally able to focus more easily and more thoroughly on the attentional demands set by one sentence than by a stretch of discourse. This contrasts with the overall more taxing challenges of retrieving temporal relations from extended discourse, which may often link, for example, morphological means of expressing temporality (in one sentence) with lexical means of expressing temporality (in other sentences).98 In other words, if we assume that there are three means of expressing temporality (pragmatic, lexical, morphological), we could state that the expression of L2 temporality at a sentential level generally requires the use of these means in closer proximity than is the case in the expression of L2 temporality at a suprasentential level, which may see the means of L2 temporality in more complex distributions. Learning the past/present perfect distinction in predominantly closed environments (e.g., controlled instructional contexts requiring focus on accuracy) may be equated

98 Factors such as lexical density, sentence length, etc. may, of course, also play a role in determining the complexity at a sentential level.
with the learning of less challenging skills than learning the distinction in predominantly open environments (e.g., communicative instructional contexts), which may be equated with the learning of more challenging skills.

The open/closed distinction may also be reflected in outcome measures that are used for testing purposes. Translation task types (e.g., the translation of an entire sentence) require L2 learners to use skills which may be considered predominantly open skills since the environments in which the skills appear present L2 learners with less stability than, for example, environments in which focused grammaticality judgement task types appear (e.g., judging the grammaticality of a contextualized temporal verb form). The less stable environment, that is, the more open environment, may be the result of a combination of challenges that L2 learners face: grammatical, lexical, stylistic, etc.

In sum, maybe we should not conceive of skill specificity as a dichotomous construct but rather as a continuum, which may present itself as a result of a various contributing factors. Increased research into skill specificity and closed skills and open skills in SLA contexts could indeed shed light on the possible fluid nature of skill specificity. However, since this was not the focus of the research in this dissertation, the possibility provided above should be regarded as tentative and requires further exploration.

7.4 Research strengths and limitations
The experimental research that was carried out used existing design-related and methodological features but combined them in such a way that the resultant experimental setup was unique and showed several strengths.

The first strength worthy of attention is the thoroughness and consistency of the design and methodology applied to all three of the studies. The complex combination of variables (one dependent variable, four independent variables) resulted in both hierarchical and non-hierarchical relationships, which not only had to be made available in an accessible, user-friendly, computer-guided environment for the participants but also had to be accessible for the retrieval of relevant data. In this respect, the thoroughness of the computer programming was extremely conducive to data retrieval and data analysis. Even though changes were made during the progression from one study to the next, the computer programming remained stable and reliant throughout.

The second strength of the experimental research carried out is its relative novelty. Studies into the acquisition of L2 temporality in English make up a large part of SLA research into L2 temporality. However, even though specific features of the complex past/present perfect distinction in English have been the focus of investigation in past SLA studies (e.g., Bardovi-Harlig, 1997; Collins, 2002), studies into L2 temporality in English (in general and the past/present distinction in particular) and into features such as input practice, output practice, input processing and skill acquisition are few and far between. Bardovi-Harlig (2000) acknowledges that “the integration studies of input and
input processing with studies of production of temporal expression would contribute to a fuller understanding of the acquisition of temporal expression” (p. 412). She continues her discussion by addressing the merits of investigating additional areas of investigation that are not production-oriented (e.g., acceptability judgement studies) and states that “these types of studies may also lead to a better understanding of different levels of learner knowledge” (p. 413). The last decade has seen an influx in the number of studies into, for example, L2 temporality and input processing/processing instruction but the majority of those studies have not investigated L2 verb-phrase morphology in English but rather in Romance languages such as French (e.g., Laval, 2008; VanPatten & Wong, 2004), Italian (e.g., Benati, 2001, 2004) and Spanish (e.g., Cheng, 2002; Farley, 2001a, 2001b, 2004b). In this respect, the experimental research in this dissertation should be seen as a contribution to the relative lack of studies into L2 temporality in English in the fields referred to above.

The third strength is the computer-assisted language learning (CALL) environment in which the studies were carried out. Not only was the CALL environment conducive to increasing experimental control over variables (internal validity) but it also showed how the input-practice and output-practice treatments provided in the studies could be delivered in a manner which is in perfect harmony with methods of contemporary computer-assisted language teaching (ecological validity). In addition, it allowed for the provision of consistent quality feedback across the three studies and across all of the practice-based treatments. Discussions of the effectiveness of (consistent) feedback highlight the importance of this experimental feature (see, for example, Li, 2010, for a meta-analysis of the effectiveness of corrective feedback).

The fourth strength may be found in the linguistic contexts provided for the instruction of the target features in all three of the studies. Because temporality may be expressed using various means, it was essential to integrate these means of expressing temporality into authentic language use. The selected linguistic contexts provided contextualized samples of written communication (as opposed to decontextualized sentences), in which general messages were delivered. In turn, the messages were essential when focusing on the meaning- and use-related aspects of the target features under investigation. This resulted in FFI which placed the overall focus on meaning and use.

As is the case with any form of experimental research, the studies carried out for this dissertation also suffered from some limitations. The following paragraphs will highlight some of the most important limitations. In discussing the limitations, references will also be made to the theory presented in Chapters 1, 2 and 3.

The first limitation that needs to be addressed may be found in the overall research approach that was adopted. By selecting the target features, that is, the past/present distinction when used to locate bygone situations in present-day English, we were able to place the focus of interest on a highly problematic
temporal form–meaning–use mapping in English for ESL learners of all levels of proficiency. However, in so doing, the research was not only given more focus, it was also drastically reduced. It is this high degree of reductionism which many may see as a limitation in the research that was carried out. In the discussion of the grammatical category of tense in Chapters 1, 2 and 3, references were made to the complexity inherent in the past/present perfect distinction. In addition, references were also made to the broader picture of L2 temporality and to temporality in present-day English. The past/present perfect distinction is but one (problematic) feature found in the English tense system. As such, it represents only one challenge that L2 learners face when acquiring temporal FMU mappings in present-day English. In addition to the ‘internal’ complexity that the English tense system displays as a result of the number of temporal challenges that ESL learners may face, there is also the feature of ‘external’ complexity. When acquiring tense, ESL learners are also faced with challenges which extend beyond mere temporal problems. Often tense interacts with other grammatical categories (e.g., aspect, mood) and, in so doing, creates additional challenges for ESL learners, who may also see themselves faced with, for example, tempo-/aspectual challenges. The temporal web is a large one, so by focusing on just one temporal problem—as was the case in Studies 1, 2 and 3—the challenge of acquiring tense was inevitably reduced.

The second limitation is related to the complexity of the target feature under investigation and its selection as an item for FFI. In Chapter 2, a description of the complexity-related issues that ESL learners face when acquiring temporality in present-day English was provided. The overall conclusion was that the concept of temporal SLA verb-phrase complexity showed intricate relationships with a variety of internal (e.g., inherent form-related, meaning-related and use-related complexity) and external factors (e.g., L1 transfer, outcome measure complexity). The questions that could then be raised with respect to the complexity of the target features are the following: How helpful can instruction be for a target feature with attested complexity? Are the effects of instruction not inherently limited by the complexity that a target feature displays? Are certain forms of instruction maybe not optimal for complex target features? All these questions address the teachability of L2 target features. Lightbown (2004) highlights some criteria for the selection of target features as candidates for processing instruction. In her discussion of the criteria used to select language features for instruction, she states the following:

This would include those [language features] that learners acquire without apparent difficulty while they engage in interactive communicative language. At the other end of the continuum are those that learners continue to have difficulty with, because of their inability to distinguish between correct and incorrect language forms. These are features that do not ordinarily lead learners to misinterpret what they hear and read, and accuracy in producing these features leads to a more polished performance rather than to changes in the meaning. (p. 73)
The past/present perfect distinction can undoubtedly be placed on the more complex end of Lightbown’s continuum. ESL learners of all levels of proficiency experience problems with the past/present perfect and are often unable to distinguish between correct and incorrect uses of both tenses when locating bygone situations in English. Many (Dutch-speaking) ESL learners regard a grammatically incorrect sentence such as *I have spoken to him last week as grammatically correct. In addition, the use of the present perfect in sentences with past-zone adverbials does not actually lead to any misinterpretations. Such sentences may sound somewhat quirky and possibly even awkward but they do not lead to a fundamental breakdown in communication.

A third and final limitation may be found in the nature and duration of instruction provided in the three studies. If the instruction provided to the input-only groups in, for example, Studies 2 and 3 were represented as forms of IP, some IP proponents could level criticism at the nature of the practice when comparing the features of instruction in Studies 2 and 3 with the guidelines for drawing up structured input activities (SIAs). Although VanPatten (1996) clearly states that the SIA guidelines are guidelines and not maxims, discussions of replication studies by IP proponents have often stressed the nature and purpose of SIAs when contrasts of the SIAs in IP studies and in so-called replication studies have been undertaken. The input-practice treatment in Studies 2 and 3 did indeed consist of some input activities but strict adherence to the checklist of SIA guidelines was absent. An obvious example is the dichotomous distinction that is made by VanPatten and IP/PI proponents between referential and affective SIAs (see Footnote 81). The input practice that was provided in Studies 2 and 3 consisted of referential practice activities only. Another example is the SIA guideline that both oral and written input should be provided (VanPatten, 1996). This did not happen in the input-practice sessions provided in Studies 1, 2 and 3. Only written input was provided. Although there were minor departures from traditional IP treatment design, I do not believe that these departures merit the exclusion of the input practice in the studies as instances of SIA practice. Indeed, in an analysis of IP treatment design, Doughty (2004) addresses the departure from the original IP treatment design. One such departure visible in some PI studies is the move towards language-manipulation and metalinguistic activities (Doughty, 2004). As far as the duration of the instruction is concerned, one could argue that the limited duration in all three studies does not facilitate the acquisition of a complex temporal form–meaning–use mapping. Two weeks of instruction (one week of theory, one week of practice) is in effect not long but it should be borne in mind that the target feature was not one which the participants were not using at all. Most of the participants were indeed able to distinguish between the past and the present perfect when locating bygone situations on many occasions. The only problem was the consistency with which they were able to do so. In other words, acquisition was not defined as the internalization of the past/present
perfect distinction (i.e., initial emergence) but rather as increased control over the distinction. Whether short-term practice is more conducive to developing increased control over a form than to internalizing a new form is debatable and is dependent on a variety of factors (e.g., instruction intensity, target-feature frequency, target-feature complexity). Consequently, the issue of practice duration will feature in the suggestions for future research (Section 7.6).

7.5 Implications for language pedagogy

Before looking at some of the pedagogical implications of the results found in Studies 1, 2, and 3, let us return to the two fundamental questions which were highlighted in the introduction to this doctoral dissertation. Both questions encapsulate the essence of the research that has been carried out: (1) How can materials based on practice best be incorporated into settings in which explicit grammar instruction is the norm? and (2) How can materials best be selected based on the treatment effects that they may bring about?

If we look at the results in all three of the studies, it would appear that neither input processing nor skill acquisition theory correctly predicted what effectively happened to the participants’ performances in Studies 1, 2, and 3. However, should the ‘scoreboard’ show an overall 0-0 score with respect to the effects of input practice and output practice? Definitely not. An overall score of 1-1 would appear more acceptable in that both the input-practice groups and the output-practice groups in all three studies showed increases on their post-test performances. What we noticed was that the practice which was given in all three of the studies—be it input practice or output practice—in combination with explicit information about the complex L2 target feature and about processing problems led to statistically significant increases in posttest performances. I believe that this is one of the main pedagogical lessons that should be taken from the empirical research carried out. In settings in which explicit instruction is the norm, a combination of explicit information and practice—albeit short-term practice—appears to be conducive to improving L2 learners’ performances with respect to the selected target feature. How that practice should be embedded instructionally is open for discussion. The results from the findings appear to indicate that input practice and output practice do not bring about differential effects. As far as the selection of the types of practice material is concerned, I believe that contemporary learner grammars of English have already started to make noteworthy inroads into a more ‘holistic’ approach to practice by complementing traditional output-practice tasks with forms of input practice. In so doing, they have started to move away from the heavy output focus that grammars of English have shown in the past.

With respect to the second issue under investigation, that is, skill specificity, the predictions of both approaches about skill specificity were not borne out in the research that was carried out. However, we already highlighted some factors (e.g., limited practice, high degree of target-feature complexity) which may have
limited some of the effects and the presence or absence of skill specificity during the experimental sessions. We could add to that the highlighted distinction between closed and open skills and the results obtained are perhaps not as unexpected as initially observed. Further research into this line of investigation is required to provide pedagogical suggestions for improving the effects of explicit (practice-based) instruction.

ESL teachers may want their learners to understand individual target features from a metalinguistic point of view and to focus on, for example, problematic features from a contrastive point of view. However, the overall aim of acquisition cannot simply be the internalization of new target features in any shape or form. Increased control—of both a receptive and productive nature—over target features in a variety of contexts is what many (intermediate to advanced) ESL learners are generally looking for in English languages classes which are part of a larger language programme aimed at communication in the broadest sense of the word. Explicit information about the target feature and processing problems—in addition to practice of both a receptive and productive nature—will best prepare ESL learners in their challenge of mastering (complex) L2 target features in communicative contexts. However, the phenomenon of language skill learning needs to be further investigated with more nuanced approaches such as the distinction between closed and open skills.

7.6 Suggestions for further research

Taking into account the information above, I would like to provide in this section some suggestions for possible further research. Of course, these are merely a few suggestions and the constellation of factors discussed in this dissertation provides—without any doubt—fertile ground for additional promising avenues of investigation. The distinction between closed and open skills referred to above is definitely one line of investigation which needs to be pursued but since this feature has already been highlighted above, I will focus on three other issues.

Focusing on the acquisition of one specific target feature often brings with it a degree of reductionism, which was already discussed above. In turn, such a reductionist approach may lead to an incomplete or even distorted picture in the broader context of second language acquisition. Acquiring L2 temporality is a prime example in this respect since it may be analysed into smaller units (e.g., tenses) yet covers at the same time a variety of target features which all interact in extended discourse (e.g., adverbials, relative tenses). A reductionist approach was applied in the three experimental studies, which all focused on the problematic past/present perfect distinction when used to locate bygone situations in present-day English. Future research might do well to address the issue of reductionism by focusing not on one target feature in isolation but on either one problematic target feature in its relationships with less problematic target features or on several problematic target features simultaneously. Taxing L2
learners’ attentional resources with too much complexity may, at first glance, appear counterproductive. However, depending on the sequence, rate and nature of increased complexity, the incremental introduction to temporal complexity may be exactly what L2 learners need to get to grips with attentional challenges involved in acquiring L2 temporality in general and tense in particular. For example, an interesting way to introduce ESL learners to incremental complexity would be to adopt a form-oriented approach to temporality by targeting one problematic temporal form (e.g., the present perfect) and to contrast problematic uses of that temporal form. In the case of the present perfect in English, one could focus on the incorrect overgeneralization of the present perfect in past-zone contexts as in *I have seen him two days ago (see past/present distinction in Chapter 1) and on the incorrect undergeneralization of the same temporal form in other contexts (e.g., *I am here for two hours already). Both of the examples have in common problematic uses of the same temporal form, the present perfect. By addressing these problems contrastively and both intralinguistically and crosslinguistically, L2 learners may be able to focus their attentional resources on exactly those problematic uses of the present perfect and, in the long run, bring about changes in their interlanguages.

Another suggestion for future research addresses the issue of prolonged practice. Studies 1, 2 and 3 all contained different forms of practice but what they had in common was a relatively limited time of practice. The road to the automatic use of L2 target features is often a long one and this may be even more so with highly complex L2 target features. In all three of the studies, there were increases in performance but perhaps the right combination of type of instruction and practice duration may be more conducive to stabilizing L2 learners’ interlanguages with respect to accurate and meaningful uses of complex temporal FMU mappings. In addition, increased practice may also have an effect on the durability of any treatment effects although this is not a given (see Chapter 3). Prolonged practice, however, is definitely a line of research which merits further investigation.

A third and final suggestion for future research addresses the nature of the practice provided. In Studies 1, 2 and 3, we first provided the L2 learners with explicit information about the complex target structure, which was followed by either input-practice or output-practice sessions. One of the reasons why a clear distinction was made between input practice and output practice was to see whether these two different forms of practice had any differential effects on acquiring the complex temporal form–meaning–use mapping under investigation. An issue that we did not address were the potential effects of a combination of input practice and output practice. Adding another treatment group to the research design and providing participants in this additional treatment group with explicit information and both forms of practice may reveal other pedagogically viable options for administering practice-based instruction. Such a combination may give L2 learners the best of both practice-based worlds,
CHAPTER 7

which, in turn, may result in higher and more durable increases in performances.

7.7 Conclusion

The research carried out in this doctoral dissertation has contributed to at least three active areas of SLA research: (1) FFI research, (2) research about L2 temporality in English and (3) research about the roles and effects of practice in the SLA process. With respect to FFI research, the research has confirmed that there is indeed a positive role for FFI in the acquisition of temporality in settings in which explicit instruction is the norm. Although Norris and Ortega (2000) have shown in their research synthesis and meta-analysis that FFI—including explicit FFI—may be conducive to SLA, not all scholars believe in the potential of explicit FFI in the SLA process. The results from all three studies show consistent increases in the control over the past/present perfect distinction (when used to locate bygone situations) as a result of explicit FFI instruction.

With respect to the acquisition of L2 temporality, the research in this dissertation has shown that by using a well-established approach to investigating the acquisition of L2 temporality—the meaning-oriented approach—a problematic temporal form-meaning-use mapping may be targeted and selected for instruction. By taking into account complexity-related features of the targeted form-meaning-use mapping, more effective ways of acquiring and teaching L2 temporality can be investigated. Important in this respect is that temporality should not just be approached in its morphological instantiation alone but also in its intricate relationships with, for example, lexical instantiation of temporality.

Last but not least, the research has tried to investigate predictions about practice using two approaches to SLA: (1) input processing (IP)/processing instruction (PI) and (2) skill acquisition theory. Although the predictions were not confirmed by the experimental research that was carried out, we did observe performance increases as a result of practice-based instruction in combination with the provision of explicit information about the target feature under investigation and processing problems. Additional research into the roles and effects of practice on the acquisition of L2 temporality in English is required to advance promising lines of investigation. Such research could include the following suggestions: practice duration as an independent variable, combining input practice and output practice, skill specificity in closed and open environments, etc.

Practice may not make perfect but as long as it helps L2 learners on their way to form-related, meaning-related and use-related linguistic perfection, researchers and L2 teachers should do all that they can to help provide insights into practice-based instruction to create instances of perfect practice and, in so doing, optimize the SLA process.
REFERENCES


REFERENCES


Doughty, C. J. (2004). Commentary: When PI is focus on form it is very, very good, but when it is focus on forms... In B. VanPatten (Ed.), Processing instruction: Theory, research, and commentary (pp. 257–270). Mahwah, NJ: Lawrence Erlbaum.
REFERENCES

REFERENCES


REFERENCES


REFERENCES


Kuiken, M. Pierrard & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning and teaching* (pp. 210–236). Brussels: Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.


REFERENCES


APPENDIX

The following appendix contains some example slides that were provided to the participants during the treatment theory sessions in all three studies. This appendix does not contain all of the slides but simply a selection. The order of the slides represents the order in which they appeared during the sessions.

The cover term SITUATION will be used here to refer to anything that can be expressed using a verb:

- ACTIONS
  (01) John has dug a hole.

- EVENTS
  (02) The bomb exploded.

- PROCESSES
  (03) Government pressure has increased dramatically.

- STATES
  (04) He was extremely polite.
The cover term **BYGONE SITUATION** will be used here to refer to *any situation* (action, event, process or state) which can be located *before* the present time (‘now’). As is the case with any other situation, bygone situations may refer to a specific point (red circle) or to a period of time (red line) before the present time (‘now’):

---

**BYGONE SITUATIONS** (before ‘now’)

(05) *John has dug a hole.*

(06) *We talked to her for almost an hour.*

(07) *We left after the film had finished.*

---

A situation is considered a **NON-BYGONE SITUATION** if it in any way includes the present time and/or future time (green circle or line):

---

**NON-BYGONE SITUATIONS**

(08) *Jayden is being extremely rude.* (includes the present)

(09) *Next year, they will have known each other 30 years.* (includes the past, present and future)

(10) *Susan will not be here in the summer.* (includes the future)
OBSERVATIONS:

a) Both the PAST tense (I worked) and the PRESENT PERFECT tense (I have worked) can be used in English to refer to bygone situations but they are very seldom interchangeable.

b) The choice between PAST and PRESENT PERFECT is one which either has to be made by writers or speakers or may be forced upon the them (by the context, see examples later on).

c) The general idea is that the PAST tense focuses on ‘then’ (before the present), whereas the PRESENT PERFECT tense focuses on ‘now’ (the present).

d) Often, the deciding factor when choosing between the PAST tense and the PRESENT PERFECT tense is the context in which the verb is used (see observation (b) above and examples later on).

e) When a past context is provided, the PAST tense must be used in English. This is not necessarily the case in Dutch, which often uses the VOLTOOID TEGENWOORDIGE TIJD (vtt, Ik heb gewerkt) (see examples later on).

When a past context is provided (for example, by means of past-time indicators (in red below)), the PAST tense must be used in English (in blue below). As already mentioned, this is not necessarily the case in Dutch, which often uses the VOLTOOID TEGENWOORDIGE TIJD (vtt) (in green below):

COMPARE

(11) I saw him yesterday. (PAST)
(12) Ik heb hem gisteren gezien. (vtt)
(13) Two years ago we drove to Germany for the first time. (PAST)
(14) Twee jaar geleden zijn we voor de eerste keer naar Duitsland gereden. (vtt)
(15) As a child I never really did that. (PAST)
(16) Als kind heb ik dat nooit echt gedaan. (vtt) (said by someone who is now an adult)
(17) On 11 September 2001 many people lost their lives in New York. (PAST)
Different kinds of context can be distinguished and for the **PAST** vs **PRESENT PERFECT** discussion the following two contexts will be investigated:

a) **LINGUISTIC CONTEXT** (= use of (explicit) past-time indicators but not necessarily in the same clause/sentence as the past verb form!)

(19) *I spoke* to my boyfriend yesterday.
(20) *Siobhan has waited/waited* for hours. (difference in meaning!)

b) **PRAGMATIC CONTEXT** (= general/specific world knowledge)

(21) *The Romans founded* the city of York.
(22) *Luciano Pavarotti was* a famous opera singer.

---

a) **LINGUISTIC CONTEXT** (= use of (explicit) past-time indicators)

Past-time indicators can take on several forms and we will look at two of the most common past-time indicators: (i) **TIME ADVERBIALS** and (ii) **PLACE ADVERBIALS**. Often, the past-time indicators are mentioned explicitly in the same sentence in which the conjugated verb appears but in stretches of discourse/texts this may not always be the case. In such cases, the past-time indicators may be present elsewhere in the text (see examples later on).

(i) **TIME ADVERBIALS** (referring to a definite point or a period of time in the past)

e.g., *yesterday, two days ago, last month, in 1989*

(23) *She saw him a couple of minutes ago.* (PAST)
(24) *Ze heeft hem enkele minuten geleden gezien.* (vtt)
(25) *She started taking swimming classes in June and it is now September.* (PAST)
(26) *Ze is in juni met zwemlessen begonnen en het is nu september.* (vtt)
(i) **PLACE ADVERBIALS** (‘when I/you/he was in . . .’)

e.g., in Helsinki, during my childhood, whilst travelling through the States

(27) *In Helsinki, they went straight to the hotel.* (PAST)
(28) *In Helsinki zijn ze onmiddellijk naar het hotel gegaan.* (vtt)
(Both sentences said after having returned from Helsinki)

(29) *When I was in New York I met many interesting people.* (PAST)
(30) *Toen ik in New York was heb ik veel interessante mensen ontmoet.* (vtt)
(Both sentences said after having returned from New York)

Often, the **LINGUISTIC CONTEXT** (the adverbials) can be found in the same sentence in which the conjugated the verb is located (see also previous example sentences) but this is not necessarily the case in (longer) texts. In the following text, the past-time indicators (*1901-1971, all his life*) is located in both the first sentence and the third sentence, whereas the conjugated verb (*struggled*) is located only in the third sentence. Once again, watch out for Dutch, which does not necessarily require the **PAST** tense in such contexts.

Louis Armstrong (*1901-1971*) was één van de grootste jazztrompettisten van de 20ste eeuw, maar hij eindigde zijn carrière toch als zanger. Lippen zijn niet geschikt om dag in, dag uit tegen het mondstuk van een trompet te trillen. Armstrong heeft zijn leven lang met dit gegeven geworsteld. Al op het hoogtepunt van zijn loopbaan, in de jaren ’30 van de twintigste eeuw, waren zijn lippen kapot geblazen door zijn energieke techniek.

Louis Armstrong (*1901-1971*) was one of the greatest jazz trumpet players of the 20th century, but he ended his career though as a singer. Lips are not suited to vibrate against the mouthpiece of a trumpet day in and day out. Armstrong struggled with this fact all his life. Even at the height of his career, in the 1930s, his lips had been ruined as a result of his energetic blowing technique.
Another example can be found in the text below, where the past-time indicator (in 1997 and 1998) is located in the third sentence, whereas the conjugated verb (met) is located in the fifth sentence. Once again, watch out for Dutch, which does not necessarily require the PAST tense in such contexts and often uses the vtt:

Varkenspest wordt veroorzaakt door een virus. De meeste besmette dieren bezwijken aan de ziekte. Tijdens de uitbraak van de varkenspest in 1997 en 1998 werden in Nederland ongeveer 13 miljoen varkens geruimd. Het merendeel preventief, dus zonder dat ze ziek waren. Dit heeft op groeiend verzet gestuit van de bevolking, omdat vaccineren ook mogelijk was. Geënte varkens zijn uitstekend tegen de ziekte beschermd. Alleen zijn ze een probleem voor de export.

Swine fever is caused by a virus. Most of the infected animals succumb to the disease. During the outbreak of swine fever in 1997 and 1998 some 13 million pigs were culled in the Netherlands. Most of them as part of a precautionary measure, in other words, without them being sick. This measure met with growing opposition from the population since vaccination was also an option. Vaccinated pigs enjoy perfect protection from the disease. The only problem then is exporting them.

(ii) **PRAGMATIC CONTEXT** (= world knowledge)

Very often the sentence/text in which the conjugated verb appears contains no (explicit) past-time indicators (e.g., adverbials) but the language user will be forced to rely on his/her general or specific knowledge of world events.

(31) *Luciano Pavarotti was a famous opera singer.* (PAST)
    (who died on 6 September 2007)

(32) *The German Democratic Republic was a communist state.* (PAST)
    (which ceased to exist after the fall of the Berlin Wall in 1989)
SUMMARY:

a) Both the **PAST** tense and the **PRESENT PERFECT** tense can be used in English to refer to bygone situations but they are very seldom interchangeable.

b) The choice between **PAST** and **PRESENT PERFECT** is one which either has to be made by the writer/speaker (focus on ‘now’ or ‘then’) or which may be forced upon the writer/speaker (by the context).

c) The general idea is that the **PAST** tense focuses on ‘then’ (before the present time), whereas the **PRESENT PERFECT** tense focuses on ‘now’ (the present time).

If there is no explicit context available, the **PRESENT PERFECT** tense can be used in English to refer to bygone situations. In such cases, the writer/speaker referring to the bygone situation is somehow concerned with the ‘now’ rather than with ‘then’. *(red arrow)*

(33) *John has broken his leg.* (**PRESENT PERFECT**)

(The writer/speaker may be telling you this because someone may have asked how John is doing right now and you as the writer/speaker want to focus on his situation now.)
If there is no context available, the **PAST** tense can be used in English to refer to bygone situations. In such cases, the writer/speaker is somehow concerned with the ‘then’ of the bygone situation. *(no red arrow)*

(34) **John broke his leg** (PAST)
    (The writer/speaker is thinking of what happened to John at one point in the past, for example, when he was on holiday a week ago, and not about John’s situation right now.)

If there is a context available but the indicator of time can be interpreted in various ways, the writer/speaker must decide on what has to be communicated and can do so by selecting either the **PRESENT PERFECT** or **PAST**. *(red arrow)*

(37) **Susan has lived here for twenty years?**
    (The writer/speaker may be saying this at a time when Susan is still living ‘here’ and so no break between the past and the present has been established. This must be expressed by using the **PRESENT PERFECT** tense.)
If there is a context available but the indicator of time can be interpreted in various ways, the writer/speaker must decide on what has to be communicated and can do so by selecting either the **PRESENT PERFECT** or **PAST**.

(38) *Susan lived here for twenty years?*  
(The writer/speaker may be saying this at a time when Susan is no longer living ‘here’ and so a break between the past and the present has been established. This must be expressed by using the **PAST** tense.)

---

Read the text below and try to understand the overall message of the text. Do not worry too much about the meanings of individual words right now. Simply go through the text and then read the answer to the question about the text carefully bearing in mind the 3-step summary chart that you saw earlier on.

At a speed of 360 kilometres per hour (approximately 225 mph) the French Automotrice Grande Vitesse (AGV) is said to become one of the fastest trains in the world. Recently, the French train builder Alstom presented the train as the successor to the TGV, which is manufactured by the same company and can travel at a speed of ‘only’ 330 kilometres per hour (approximately 205 mph). The AGV has its engine spread over the train carriages. Just like the German ICE train, the French AGV train has its engine built into the flooring of the train carriages. As a result, the first and last carriages no longer need to be dedicated to housing engine rooms.

However, whether train-spotters will be able to admire the new train on the French tracks remains to be seen though. *Tot nu toe heeft de Franse treinmaatschappij nog geen bestelling geplaatst.*

**Question:**
*Translate the sentence in red in the text above into standard formal British English.*

**Possible answer:**
*Up until now the French railway company has not placed an order yet.*
Het concept oefening (practice in het Engels) is voor veel taalleerders en taalonderwijzers een noodzakelijke voorwaarde voor het leren van een vreemde taal. Dit concept is echter tot nu toe relatief zelden besproken op basis van bestaande theorieën. In het verleden werd het Engelse tijdsysteem in les-verband grotendeels aan de hand van driloefeningen geoefend, die gericht waren op output en waarvoor de taalleerders grammaticaal correcte vormen dienden te produceren. Deze aanpak wordt ook gebruikt in hedendaagse taallessen en is zeker bruikbaar gebleven op zijn eigen manier, maar het output-gebaseerde oefenen (output practice) wordt in het hedendaagse lesmateriaal gecompleteerd met andere oefeningsvormen. Zo is het lesmateriaal dat is opgesteld voor expliciete instructie over het Engelse tijdsysteem veranderd door bijvoorbeeld de integratie van input-gebaseerd oefenmateriaal (input practice), dat van talisten vraagt dat ze oefenen met input zonder vormen te produceren.

Dit doctoraal proefschrift behandelt het concept practice door de effecten te onderzoeken van input practice en output practice op de verwerving van het Engelse tijdsysteem door Nederlandstaligen. In hoofdstuk 1, 2 en 3 wordt theoretische achtergrondinformatie gegeven over de doelstructuren die geselecteerd werden voor het onderzoek. Vervolgens worden in hoofdstuk 4, 5 en 6 aan de hand van drie computergestuurde leerexperimenten twee aanpakken tot taalverwerving (T2-verwerving) vergeleken, die verschillende rollen toekennen aan practice. Tenslotte worden de belangrijkste punten van deze hoofdstukken samengebracht in een algemene bespreking van de experimentele gegevens in hoofdstuk 7.

Hoofdstuk 1, Tense and linguistics, bestaat uit een linguïstische inleiding van de onderzochte doelstructuren. Deze doelstructuren zijn de onvoltooid verleden tijd (past) en de voltooid tegenwoordige tijd (present perfect), die worden gebruikt om zogenaamde bygone situations te beschrijven. Declerck (1991, 2003, 2006) defineert bygone situations als situaties die zich hebben voorgedaan vóór het temporele referentiepunt. Het doel van dit eerste hoofdstuk is drievoudig: (1) de grammaticale categorie tempus definiëren en contextualiseren, (2) tempus conceptualiseren door gebruik te maken van Declercks descriptieve theorie van het Engelse tijdsysteem en (3) de verschillende manieren vergelijken waarop de past en de present perfect worden gebruikt in het Engels, met aandacht voor de verschillen tussen deze twee typologisch sterk verwante talen. Wanneer men bestaande definities van tempus met elkaar vergelijkt, is het duidelijk dat er steeds specifieke items terugkeren waarover onderzoekers het eens of oneens zijn. Linguïsten zijn het er bijvoorbeeld over eens dat er een zeker verband bestaat tussen tempus en tijd. Er is echter geen consensus over de morfologische aard van tempus. Sommige linguïsten beweren dat tempus uitsluitend synthetisch kan worden uitgedrukt (m.a.w. door gebruik te maken van gebon-
den morfemen) en dat analytische vormen in een tijdensysteem (bv. vormen zoals de Engelse voltooide tegenwoordige tijd en de toekomstige tijd met will, die gevormd worden door combinaties van vrije morfemen) niet als tijdsvoor-}
men kunnen worden aanvaard op basis van hun morfologisch analytische aard. Na de bespreking van deze onderwerpen wordt bewust de beslissing genomen om Declercks descriptieve theorie van het Engelse tijdensysteem te volgen. Deze theorie aanvaardt dat bepaalde analytische werkwoordsvormen (bv. *present perfect*) daadwerkelijk tijden zijn, aangezien ze een temporale structuur hebben die uniek is en niet in andere werkwoordsvormen kunnen worden teruggevonden.

Hoofdstuk 2, *Tense and complexity*, vormt een brug tussen de puur linguïstische aspecten van tempuscomplexiteit en de uitdagingen waarmee taalleerders van het Engels, waaronder ook Nederlandstalige taalleerders, worden geconfronteerd bij het verwerven van het Engelse tijdensteem. De essentie van dit hoofdstuk bestaat uit het conceptualiseren van de term *temporal SLA verb-phrase complexity*, meer specifiek de complexiteit bij het verwerven van het onderzochte *past/present perfect*-onderscheid. Een uitvoerige definitie van *SLA verb-phrase complexity* wordt aan het begin van het hoofdstuk gegeven. Vervolgens worden belangrijke aspecten van de definitie gedetailleerd toegelicht met verwijzingen naar linguïstiek en tweede-taalverwerving. De hoofdidee waarop de definitie is gebaseerd, is dat de complexiteit die onderzocht wordt een concept is dat uit verschillende facetten bestaat en dat inzichten van verschillende subdisciplines vereist zijn om deze complexiteit te analyseren. Tegenovergestelde concepten zoals absolute versus relatieve linguïstische complexiteit, globale versus lokale linguïstische complexiteit, complexiteit als een afhankelijk of onafhankelijke variabele worden in dit hoofdstuk toegelicht aan de hand van concrete tempusgerelateerde voorbeelden van de werkwoordelijke groep in het Engels. De twee onderwerpen die in alle delen van dit hoofdstuk terugkomen, zijn de definitie van complexiteit die aan het begin van het hoofdstuk wordt gegeven en De Keysers (2005) bespreking van wat het leren van T2-grammatica moeilijk maakt. De gedetailleerde bespreking van temporele *SLA verb-phrase complexity* – meer bepaald de complexiteit van het onderscheid tussen de *past* en de *present perfect* in gevallen waar ze worden gebruikt voor de beschrijving van *bygone situations* – houdt rekening met vorm-, betekenis- en gebruiksgerelateerde complexiteit. Bovendien wordt de complexiteit van het verbinden van vorm, betekenis en gebruik, ook wel *mapping complexity* genoemd, besproken aangezien de verwerving van vorm-betekenis-gebruik-mappings het eigenlijke doel van T2-verwerving is. De conclusie is dat *temporal SLA verb-phrase complexity* op een ingewikkelde wijze verstrengeld is met een hele reeks factoren die betrekking hebben op de T2-doelstructuren en op het T2-verwervingsproces zelf. Een onderscheid kan worden gemaakt tussen universele en specifieke factoren die de complexiteit van (grammaticale) doelstructuren bepalen: universele factoren (bv. vorm-, betekenis-, gebruiksgerelateerde complexiteit) zijn aspecten die
door de meeste, al dan niet door alle, taalleerders van het Engels als complex worden ervaren, terwijl specifieke factoren aspecten zijn die typerend zijn voor bepaalde taalleerders en/of onderzoeksontwerpen (bv. T1-invloed, outcome measures). Ondanks het feit dat een onderscheid wordt gemaakt tussen deze factoren, kan er een wisselwerking plaatsvinden tussen beide soorten, wat het onderscheid van de twee soms ingewikkeld maakt.

In hoofdstuk 3, *Tense and instructed SLA*, ligt het accent op de grammaticale categorie tempus in de context van gestuurde T2-verwerving. Het eerste gedeelte van hoofdstuk 3 is gewijd aan de studie van temporaliteit in T2-verwervingsgerelateerd onderzoek, dat incidenteel is ontstaan maar zich heeft ontwikkeld tot een goed gestructureerd, systematisch en methodologisch ontwikkeld onderzoeksgebied met verschillende vormen van onderzoek. Over het algemeen worden twee benaderingen gebruikt om T2-temporaliteit te onderzoeken: (1) een vormgeoriënteerde benadering en (2) een betekenisgeoriënteerde benadering. De lezer krijgt een inleidend overzicht van beide aanpakken, die afzonderlijk en contrastief worden besproken. Bovendien worden de drie experimenten die werden uitgevoerd voor het experimentele onderzoek van dit proefschrift geplaatst in het kader van de twee benaderingen tot T2-temporaliteit. In het tweede gedeelte van hoofdstuk 3 wordt de focus geplaatst op gestuurde T2-verwerving, meer specifiek op *form-focused instruction* (FFI), instructie die gericht is op vorm, en op twee bekende aanpakken tot T2-verwerving die *practice-gebaseerde instructie* integreren als een wezenlijk onderdeel van het T2-verwervingsproces: (1) *input processing* en (2) *skill acquisition theory*. Een beschrijving van de kenmerken van deze twee aanpakken en een bespreking van de rol van *practice* in beide benaderingen zorgen ervoor dat de lezer de noodzakelijke achtergrondinformatie krijgt om het experimentele onderzoek, dat beschreven wordt in hoofdstukken 4, 5 en 6, te beoordelen. De rol van *practice* is in beide aanpakken verschillend en een van de grote verschillen is de aanwezigheid of afwezigheid van *practice*-gerelateerde vaardigheidsspecificiteit (*skill specificity*). *Input processing* beschouwt de verwerving van taalvaardigheden niet als vaardigheidsspecifiek. Dit betekent dat de focus van deze benadering (*input* en *input practice*) de T2-verwerver zal helpen bij het verwerven van zowel receptieve als productieve vaardigheden. *Skill acquisition theory* daarentegen beschouwt de verwerving van taalvaardigheden wel als vaardigheidsspecifiek. Vanuit het standpunt van *skill acquisition theory* zal *input practice* de T2-verwerver helpen bij het verwerven van receptieve vaardigheden en *output practice* bij het verwerven van productieve vaardigheden. Dit grote verschil wordt aangekaart in de onderzoeksvragen in hoofdstuk 4, 5 en 6.

Hoofdstuk 4, *Study 1*, is het eerste van drie experimentele hoofdstukken. *Study 1* werd opgesteld met *treatment* als *between-groups*-variabele waarin drie niveaus (input only, input practice, output practice) werden gedifferentieerd. Er namen 70 studenten Engels deel aan het onderzoek. We gingen ervan uit dat expliciete FFI daadwerkelijk een effect zou hebben op de prestaties van de deelnemers bij
een tempusgerelateerde test. Bijgevolg lag de nadruk niet op de vraag of FFI wel of geen effect zou hebben, maar op de vraag of er verschillende effecten konden worden vastgesteld met betrekking tot de drie vormen van instructie die werden gegeven. Bovendien waren we ook geïnteresseerd in de vraag of er enige vorm van interactie was tussen de vorm van instructie en de soorten effecten (skill specificity). Gebruik makend van hoofdzakelijk kwantitatieve methoden van dataverzameling, werden de volgende drie within-groups-variabelen onderscheiden: (1) tijd (pretest, posttest), (2) modaliteit (receptief, productief) en (3) outcome measure (grammaticality judgement, selected response, constrained constructed response, translation). De drie vormen van instructie werden vervolgens geoperationaliseerd. Uit de analyse van de data bleek een kleine toename van pretest-naar-posttestprestatie (ongeveer 1,5 op een maximumscore van 24), maar dit effect van tijd was niet het gevolg van treatment, modaliteit of outcome measure. Met andere woorden, de drie treatment-groepen vertoonden alle drie een geringe vooruitgang wat hun prestatie op de pretest en posttest betrof. Het was echter onmogelijk om vast te stellen of deze toename het gevolg was van treatment, aangezien we beslist hadden geen controlegroep te integreren in het onderzoek. Ten gevolge van deze resultaten werden vier veranderingen voor het onderzoeksdesign en de methodologie geformuleerd. De bedoeling hiervan was een tweede, verbeterd experiment te kunnen uitvoeren. Ten eerste werd het aantal treatments verminderd van drie (input only, input practice, output practice) naar twee (input practice, output practice). Ten tweede werden de input practice en de output practice waaraan de deelnemers werden blootgesteld, toegespitst op de outcome measure van de pretest en posttest. Ten derde werd het aantal testitems verhoogd met acht items, wat het totaal aantal testitems op 32 bracht. En ten vierde werden uitsluitend tempusgerelateerde vragen gesteld op de tests.

Hoofdstuk 5, Study 2, rapporteert over het tweede experiment, dat één jaar na het eerste werd uitgevoerd. Er werd gewerkt met compleet nieuwe deelnemers, maar de onderzoeksvragen waren in wezen dezelfde als de onderzoeksvragen geformuleerd voor Study 1. Na de implementatie van de vier veranderingen vermeld in de vorige alinea, werden de data van de 71 deelnemers onderworpen aan kwantitatieve analyses. Deze keer brachten de analyses een grote vooruitgang met betrekking tot pretest- en posttestprestaties aan het licht (ongeveer 4,5 op een maximumscore van 32). Alhoewel een grote toename werd vastgesteld, was het nog steeds onmogelijk om vast te stellen of deze toename het gevolg was van treatment. We hadden immers beslist om ook nu geen controlegroep op te nemen in het onderzoek. Aangezien de toename deze keer groot was, werd de beslissing genomen een derde experiment uit te voeren.

Hoofdstuk 6, Study 3, is gewijd aan het derde en laatste experiment dat is uitgevoerd. Study 3 was in principe een herhaling van Study 2 maar er werden twee zaken toegevoegd: (1) een controlegroep om het effect van treatment te onderzoeken en (2) een uitgestelde posttest om de duurzaamheid van mogelijke effecten te bestuderen. Study 3 is één jaar na Study 2 uitgevoerd, met 75 deelne-
De onderzoeksvragen waren dezelfde als de onderzoeksvragen geformuleerd voor Study 1 en Study 2. Door de toevoeging van een uitgestelde posttest hebben we beslist een vierde onderzoeksvraag te formuleren om zo de duurzaamheid van mogelijke treatment-effecten te onderzoeken. Het aantal treatments uit Study 2 werd behouden (input practice, output practice). De controlegroep kreeg echter ook een vorm van treatment (die geen verband hield met tempus, maar met de grammaticale categorie getal). Aangezien deze vorm van treatment niets met tempus had te maken, was de eigenlijke aard van treatment in de controlegroep niet belangrijk voor de onderzoeksvragen. Ook in dit experiment werd weer een grote toename van pretest- naar posttestprestatie gevonden (ongeveer 3,5 op een maximumscore van 32). De analyse toonde significante toenames in testprestatie (van pretest naar onmiddellijke posttest) voor zowel de input practice- als de output practice-groep. Er werden echter geen significante toenames waargenomen voor de controlegroep. Een vergelijking van de data van de input practice- en de output practice-groep wees niet op significante verschillen tussen beide groepen. Uit de analyse van de data van de uitgestelde posttest bleek dat het treatment-effect duurzaam was ondanks een kleine afname met betrekking tot de prestaties op de uitgestelde posttest.

In hoofdstuk 7, General discussion, wordt een overzicht gegeven van de conclusies van alle drie experimenten. Hoofdstuk 7 brengt dus de theoretische en experimentele hoofdstukken samen door zowel een samenvatting als een meer gedetailleerde bespreking van de conclusies van de drie experimenten te bezorgen. De expliciete FFI die aan de deelnemers werd gegeven, bracht daadwerkelijk significante toenames tot stand in de posttestprestaties van de deelnemers en dit met betrekking tot de complexe, temporele vorm-betekenis-gebruik-mappings die onderzocht werden. Ondanks het feit dat de toename in Study 1 relatief klein was, leidde practice die was toegespitst op de outcome measures tot grote toenames in Study 2 en Study 3. Bovendien bleken in Study 3 de toenames in testprestaties van duurzame aard te zijn wanneer we de prestaties op de pretest, onmiddellijke posttest en uitgestelde posttest vergeleken. In de drie experimenten werden geen verschillende effecten vastgesteld tussen deelnemers die input practice hadden ontvangen en deelnemers die output practice hadden ontvangen. Daarenboven werden er geen interacties gevonden tussen de vorm van practice (input practice, output practice) en de modaliteit van de outcome measures (receptief, productief). Een van de hoofdzaken die onderzocht werden, was eventuele skill specificity bij het verwerven van taalvaardigheden. Volgens voorspellingen van input processing is de verwerving van T2-doelstructuren niet vaardigheidsspecifiek. Dit staat in contrast met de voorspellingen van skill acquisition theory, waar overgang van de ene ontwikkelingsfase naar de volgende wel vaardigheidsspecifiek is. Met andere woorden, input practice leidt tot toenames voor receptieve vaardigheden (outcome measures) terwijl output practice leidt tot toenames voor productieve outcome measures. De voorspellingen die gebaseerd zijn op practice (volgens beide aanpakken) werden niet bevestigd in de drie expe-
Een mogelijke verklaring voor deze resultaten zou kunnen gevonden worden als een meer genuanceerde conceptualisatie van *skill specificity* wordt gebruikt. In plaats van te focussen op de twee uiteinden van het continuüm voor *skill specificity*, moeten we misschien kijken naar mogelijke plaatsen tussen vaardigheidsspecifiek en niet-vaardigheidsspecifiek. Interactie met andere factoren (bv. de open/gesloten aard van vaardigheden) zou kunnen leiden tot een nauwkeuriger afgestemd beeld van hoe taalvaardigheden in bepaalde contexten worden geleerd. Daar er geen experimenten bestaan die deze mogelijke interactie onderzoeken, werden er geen zekere claims gemaakt. Naast een gedetailleerde bespreking van de resultaten biedt hoofdstuk 7 ook een bespreking van de sterke punten en beperkingen van het experimentele onderzoek. Sterkte punten zoals designgerelateerde en methodologische grondigheid en consistentie, het innoverende karakter van de onderzoeksfocus toegepast op T2-Engels en de leeromgeving die computerondersteunend leren van taal promoot voor alle drie experimenten droegen alle bij tot de kwaliteit van het onderzoek. Er waren echter ook beperkingen in de vorm van het hoge niveau van reductionisme, de moeilijke aard van de complexe doelstructuren voor instructie, en de aard van de instructie die in de experimenten werd gegeven. Hoofdstuk 7 bespreekt ook de implicaties van de resultaten voor taalpedagogie. Het hoofdstuk wordt afgerond met voorstellen voor toekomstig onderzoek, waarbij rekening wordt gehouden met de sterktes en beperkingen die reeds beschreven zijn. De uiteindelijke conclusie is dat oefening misschien geen perfecte kunst baart, maar zolang oefening T2-verwervers helpt op weg naar vorm-, betekenis-, en gebruiksgerelateerde linguïstische perfectie, zouden T2-onderzoekers en docenten al het mogelijke moeten doen om inzichten in *practice*-gebaseerde instructie te verschaffen om zo voorbeelden van perfecte *practice* te creëren en op die manier het T2-verwervingproces te optimaliseren.
The concept of practice as a prerequisite for learning a foreign language is a given for many language learners and language teachers. However, it has received relatively little attention from a theory-based perspective. In the past, much of the instructional practice related to tense in English consisted of mainly output-focused drills, which required learners to produce grammatically correct forms. Although this approach has found its way into contemporary language teaching and has proven useful in its own ways, it has focused almost exclusively on output practice to the exclusion of other forms of practice. Contemporary materials designed for explicit instruction on tense in English have seen changes such as the addition of input-based practice, which requires learners to interact with input without producing any forms.

This doctoral dissertation addresses the concept of practice by investigating the effects of input practice and output practice on the acquisition of tense by Dutch-speaking learners of English. It does so by providing—in Chapters 1, 2 and 3—relevant, theoretical background information about the L2 target features under investigation. In Chapters 4, 5 and 6, two approaches to second language acquisition (SLA), which assign different roles to practice, are compared in three computer-controlled learning experiments. Finally, all the chapters are brought together in a general discussion of the findings in Chapter 7.

In Chapter 1, Tense and linguistics, a linguistic introduction to the L2 target features under investigation is provided. The target features under investigation are the past and the present perfect when used to locate bygone situations in present-day English. Declerck (1991, 2003, 2006) defines bygone situations as situations which have taken place before the temporal zero point. The aim of this chapter is threefold: (1) to define and contextualize the grammatical category of tense, (2) to conceptualize tense using Declerck’s descriptive theory of tense in English and (3) to compare the uses of the past/present perfection distinctions in English and Dutch, with a special focus on the differences between these two typologically closely related languages.

When existing definitions of tense are compared with each other, it is clear that recurring items of both agreement and disagreement are visible. For example, linguists are in agreement that there is some form of relationship between tense and time. By contrast, there is no consensus on the morphological nature of tense, with some linguists claiming that tense is expressed only synthetically (i.e., by means of bound morphemes), and that analytic tense forms (i.e., forms such as the present perfect and the will-future, which are produced by combining free morphemes) cannot be accepted as tense forms because of their morphologically analytic nature. After the discussion of relevant items of agreement and disagreement, a conscious decision is taken to follow Declerck’s descriptive theory of tense in English, which accepts that certain analytic verb forms (e.g.,
present perfect) are tenses in their own right since they have temporal structures which are unique and cannot be found in any other verb forms.

Chapter 2, *Tense and complexity*, builds a bridge between the purely linguistic aspects of tense complexity and the challenges that ESL learners—including Dutch-speaking ESL learners—face when trying to acquire tense in English. At the heart of this chapter is the conceptualization of the term temporal SLA verb-phrase complexity, more precisely SLA complexity related to the past/present perfect distinction under investigation. A comprehensive definition of SLA verb-phrase complexity is provided at the start of the chapter. Subsequently, relevant features of the definition are explained in detail with references to linguistics and SLA throughout the chapter. The core idea on which the definition is based is that the complexity under investigation is a multifaceted concept which requires insights from a variety of subfields. Oppositions such as absolute versus relative linguistic complexity, global versus local linguistic complexity, complexity as a dependent or independent variable are intertwined in Chapter 2 with concrete tense-related examples related to the English verb phrase. The two themes that run through all the sections of the chapter are the definition of complexity provided at the start and DeKeyser’s (2005) discussion of what makes learning L2 grammar difficult. The detailed discussion of temporal SLA verb-phrase complexity, more precisely, the complexity involved in distinguishing between the past and the present perfect to locate bygone situations, takes into account, form-related, meaning-related and use-related complexity. In addition, mapping-related complexity is discussed since the acquisition of form-meaning-use mappings is the actual goal of SLA. The overall conclusion is that temporal SLA verb-phrase complexity is intricately interwoven with a variety of factors in both the L2 target features which are selected and in the SLA process itself. A distinction may be made between universal and specific items of complexity, with universal features (e.g., form-related, meaning-related, use-related complexity) being experienced by most—if not all—ESL learners and specific features, which are characteristic of specific constellations of factors related to the language learner and/or research design (e.g., L1 transfer, outcome measures). Although a distinction between these two types of items is made, forms of interplay arise, which can make it difficult to distinguish between the two.

In Chapter 3, *Tense and instructed SLA*, the focus is placed on the grammatical category of tense in the context of instructed SLA. The first section of Chapter 3 is dedicated to the study of temporality in SLA research, which came about incidentally but has developed into a well-structured, systematic and methodologically developed area of research with various lines of investigation. Two approaches have generally been used to investigate L2 temporality: (1) a form-oriented approach and (2) a meaning-oriented approach. The reader is introduced to both approaches, which are discussed in isolation and contrastively. In addition, the three studies that were carried out for the experimental
SUMMARY IN ENGLISH

part of the doctoral dissertation are discussed in the context of the approaches to L2 temporality. In the second section of Chapter 3, the focus is placed on instructed SLA—more specifically on form-focused instruction (FFI)—and on two well-known approaches to SLA which both feature practice-based instructional setups as integral parts of the SLA process: (1) input processing and (2) skill acquisition theory. A description of the features of both approaches and a discussion of the roles of practice in both approaches provide the reader with the necessary background to be able to evaluate the experimental research carried out in Chapters 4, 5 and 6. The roles of practice in both approaches are different and one of the big differences is the presence or absence of practice-related skill specificity. Input processing does not regard the acquisition of language skills as skill-specific. This means that input practice will benefit the L2 learner when acquiring both receptive and productive skills. However, skill acquisition theory does regard the acquisition of language skills as skill-specific. This means that input practice will benefit the L2 learner when acquiring receptive skills and output practice when acquiring productive skills. This major difference is addressed in the research questions in Chapters 4, 5 and 6.

Chapter 4, Study 1, is the first of three experimental chapters. Using treatment as a between-groups variable with three levels (input only, input practice, output practice), Study 1 was drawn up, in which 70 students of English participated. We assumed that explicit FFI would indeed have an effect on the participants’ performance on a tense-related test. Consequently, the focus of interest was not on whether FFI would or would not have an effect but on whether any differential effects could be found with respect to the three forms of instructional treatment provided. In addition, we were also interested in whether there was any interaction between the type of treatment and the types of effects (skill specificity). Using mainly quantitative data-collection methods, we decided on the following three within-groups variables: (1) time (pretest, posttest), (2) modality (receptive, productive) and (3) outcome measure (grammaticality judgement, selected response, constrained constructed response, translation). The three treatments were operationalized. Data analyses revealed a small increase from pretest to posttest performance (approximately 1.5 on a total score of 24) but this effect of time was not mediated by treatment, modality or outcome measure. In other words, all three treatment groups slightly improved their performance from pretest to posttest. However, it was impossible to state whether this increase resulted from any treatment since we decided not to include a no-treatment (control) group. As a result of these findings, four changes to the research design and methodology were formulated with a view to carrying out an improved second study. First, the number of treatments was reduced from three (input only, input practice, output practice) to two (input practice, output practice). Second, the input and output practice administered to the two treatments was geared towards the outcome measures used in the pretest and posttest. Third, the total number of test items was increased by
eight. This brought the total number of test items to 32. And fourth, only tense-related questions were asked on the tests.

Chapter 5, Study 2, reports on the second experiment, which was carried out one year after the first. The participants were new but the research questions were essentially the same as the ones formulated in Study 1. Having implemented the four changes formulated above, once again quantitative analyses were carried out. This time the data from 71 participants were used. Data analyses revealed a large increase from pretest to posttest performance (approximately 4.5 on a total score of 32). Although a large increase was found, it was still impossible for us to state whether this increase resulted from any treatment since we decided not to include a no-treatment (control) group. Since the increase in performance was large this time around, we decided to carry out a third experiment.

Chapter 6, Study 3, is dedicated to the third and final study that was carried out, which was carried out a year after Study 2. In essence, Study 3 was a replication of Study 2 but with the inclusion of both a no-treatment (control) group to address the effect-of-treatment claim and a delayed posttest to address the durability of possible effects. The data from 75 participants were analysed and the research questions were the same as the ones formulated in Studies 1 and 2. With the addition of a delayed posttest, we were able to add a fourth research question to address the durability of any treatment effects. In effect, the number of treatments was kept the same as in Study 2, which had two treatments (input practice, output practice). However, the no-treatment (control) group was also given a form of treatment (not related to tense but to the grammatical category of number). Since this treatment was not tense-related, the actual nature of the treatment was not important for the research questions. For Study 3, data analyses revealed once again a large increase from pretest to (immediate) posttest performance (approximately 3.5 on a total score of 32). The analyses showed significant increases in test performances (from pretest to immediate posttest) for both the input-practice group and the output-practice group. However, no significant increases were found for the no-treatment group. A comparison of the data from the input-practice and output-practice groups revealed no significant differences between both groups. In other words, no differential effects between the two treatment groups were confirmed. Analyses of the delayed posttest data revealed that the treatment effect appeared to be durable although a small decrease was observed as far as delayed-posttest performance was concerned.

In Chapter 7, General discussion, an overview of the findings from all three studies is provided. As such, Chapter 7 brings the theoretical and experimental chapters together by providing both a summary and a more detailed discussion of the findings from the three studies that were carried out. The explicit FFI provided to the participants did indeed bring about significant increases in the participants’ posttest performances with regard to the complex temporal form-
meaning-use mappings under investigation. Although the increase was relatively small in Study 1, practice geared towards the outcome measures led to large increases in Studies 2 and 3. Moreover, the increases in test performance appeared to be of a durable nature when comparing pretest, immediate-posttest and delayed-posttest performances in Study 3. In the three studies, no overall differential effects were observed between participants who received input practice and participants who received output practice. In addition, no differential interactions were found between the type of practice (input practice, output practice) and modality (receptive, productive). One of the key features under investigation was the skill specificity of acquiring language skills. According to input processing predictions, the acquisition of L2 target features is not expected to be skill-specific. By contrast, according to skill acquisition theory, progression through the developmental stages is believed to be skill-specific. In other words, input practice leads to increases for receptive skills (outcome measures) whereas output practice leads to increases for productive outcome measures. Practice-based predictions made by both approaches were not confirmed in any of the three studies. One explanation for the findings may be found if a more nuanced conceptualization of skill specificity is used. Instead of focusing on the two ends of the skill-specificity continuum, we may need to look at possible positions between skill-specific and not skill-specific. Interaction with other factors (e.g., the open/closed nature of skills and tasks) may lead to a more finetuned picture of how language skills are learnt in specific contexts. Because of the lack of studies investigating the possible interaction between these features, no definite claims have been made. In addition to discussing the findings in detail, Chapter 7 also discusses the strengths and limitations of the experimental research. Strengths such as the design-related and methodological thoroughness and consistency, the novelty of the research focus as applied to L2 English and the computer-assisted language learning (CALL) environment in which the studies were carried out all contribute to the quality of research. However, limitations are also present in the form of a high degree of reductionism, the troublesome nature of complex target features as items of instruction, the nature of the instruction provided in the studies. Chapter 7 also provides a discussion of the implications of the results for language pedagogy. By way of conclusion, suggestions for further research are provided taking into consideration the strengths and limitations of the research described above. The overall conclusion is that practice may not make perfect but as long as it helps L2 learners on their way to form-related, meaning-related and use-related linguistic perfection, L2 researchers and L2 teachers should do all that they can to help provide insights into practice-based instruction to create instances of perfect practice and, in so doing, optimize the SLA process.
CURRICULUM VITAE

Jim Ureel was born on 30 August 1976 in Antwerp (Belgium), where he grew up in a bilingual (Dutch–English) environment. After graduating from secondary school and completing a successful year as an exchange student in Germany, he decided to study languages, more specifically, applied linguistics with a specialization in translation. He chose the ‘uber-Germanic’ language combination Dutch–English–German–Danish and graduated in 1999 with a master’s degree in translation, having completed all of his coursework and a thesis on institutional multilingualism.

After working in the private sector for approximately two years, Jim Ureel applied for a position in the English section at the then Higher Institute of Translators and Interpreters in Antwerp. This is where he started his academic career in 2001 as a junior lecturer with teaching and research assignments. Since 2001, he has taught courses—in the bachelor’s and master’s programmes—including English grammar (theory and exercises), English language proficiency, English oral communication and the translation of specialized texts (e.g., academic, legal, medical, technical) from Dutch into English.

In addition to teaching English and carrying out research, Jim Ureel works as a translator (from Dutch and German into English) and as a proofreader of English. He is also a member of the editorial board of Linguistica Antverpiensia, New Series – Themes in Translation Studies (LANS – TTS), the journal of the Department of Translators and Interpreters at Artesis University College Antwerp, which investigates translation in its many oral and written manifestations.

Jim Ureel’s professional interests include, among others, (grammatical) aspect, case, English as a second/foreign language (ESL/EFL), grammar instruction, inflection, second language acquisition (SLA), tense and translation. His language interests include Basque, Danish, English, Finnish, German, Icelandic, Polish, Portuguese and Swedish.