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):
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>
<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>yes</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):77

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 [sic] 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 [sic] 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-

77 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). 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 one 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 one and 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,
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.