5 Discussion

5.1 Introduction

In this chapter, the findings reported in the previous chapter will be evaluated in relation to the theories and earlier FFI research discussed in Chapter 2. First, the research design will be evaluated in 5.2, highlighting both strengths and weaknesses. Next, in 5.3, the findings will be discussed in relation to the issues pertaining to the interface debate, SLA theory as outlined in 2.3, and previous FFI research findings (2.4). In 5.4, the implications for pedagogy will be discussed, and the book will be concluded in 5.5. Recommendations for future research will be provided along the way.

5.2 Evaluation of the research design

This study has sought to combine the advantages of laboratory research with the advantages of classroom research. In a review of SLA laboratory research, Hulstijn (1997) points out that the reason to employ laboratory research designs has been to exercise greater control over exposure to the target structures and to ensure equal conditions during the instruction: effects of instruction simply stand more chance of being spotted in tightly controlled settings. He illustrates this by quoting Carrol and Swain: "If feedback does not work in an experimental situation, it is highly unlikely that it would work elsewhere (1993, pp. 361-362)" (p. 134). Nevertheless, field research is needed to assess to what extent laboratory research findings may be generalized. After all, there is no guarantee that treatments that were found to be effective in the lab, will also work in classroom settings. In fact, this study has illustrated the limitations, as instruction was found to affect performance on decontextualized and more natural language tasks differently.

This study has controlled for a number of variables that traditionally may have caused interference in SLA classroom research. First of all, the computer was used as the means to present the instruction in order to exercise control over the input. It allowed for large numbers of students in ten different schools to receive

precisely the same instruction. The use of the computer also incorporates disadvantages, though. The instruction may be less effective, simply because the participants had to learn about structures in their L2 from written sources without any help of the teacher. Nevertheless, although the instruction was relatively brief (approximately three hours), significant explicit knowledge gains were obtained, which means that the instruction must have been effective. It may seem like a risk was taken to assume that the instruction would be effective, but the materials were tested in advance, and many studies have made use of even briefer treatments and still observed substantial progress (e.g., Day & Shapson, 2001; Jourdenais et al., 1995; Muranoi, 2000; Robinson, 1996; Salaberry, 1997). It should be pointed out, though, that although the computer was used, the experiment was not taken out of class. Being enrolled in an intensive Dutch course, the participants will have encountered the target structures in natural settings as well. In addition, there is a chance that teachers responded to questions by students, and thus that some students received explicit feedback about one of the target structures on an individual basis. In fact, one of the control group classes was excluded from the analysis for this reason.

Classroom research has seldom used true randomization to compare the effects of instruction (Norris & Ortega, 2000). Rather, intact classes have been used, and sometimes also different teachers have been used to present the instruction. In this study, the computer was used to randomly assign the participants to either of the two treatment groups. This way, both experimental conditions were equally represented by students from all participating schools. This method of randomization proved to be successful, as no significant differences were observed between the two groups on the dependent variables at the start of the experiment, on the ID variables nor on any of the background variables assessed. At the same time, this study underscores the importance of randomization: the no-treatment control group was made up of intact classes, and this group was found to differ from both treatment groups in previous exposure to the L2 (see 3.4.5). In retrospect, the control group should have been sampled in the same way as the two experimental groups, by means of the computer. However, this would have required the development of additional 'dummy' material for those students to work with.

An aspect of research design that has often featured in laboratory research is the use of (semi-) artificial languages (Hulstijn, 1997). Obviously, this is a design feature that was not adopted in this study. The focus on second language proficiency simply precluded this option. A source of variation that is introduced by using natural languages is a lack of control over the participants' history with

the target language. The participants of this study were all advanced learners, but some had needed more time than others to achieve their current level of proficiency. This variation cannot but be accepted, as it comes with the very purpose of this study. In fact, the use of impoverished linguistic environments may be one of the key factors for why lab research lacks ecological validity. The use of an artificial language allows for carefully crafting grammar structures so that they differ only with respect to the feature under investigation. However, the risk of abstracting away too much from natural languages is quite real. Hulstijn (1997) observes: "One might wonder whether some of the lab studies … investigated purely cognitive concept formation rather than (second) language learning." (p. 139).

A point of improvement would be the relative heterogeneity of the subject sample. It was already pointed out that there were differences in previous exposure to Dutch. The role of the Dutch language varied from being (one of) the language(s) used at home to being used at school only. The participants were also heterogeneous with respect to their first languages. However, some control was exercised by means of careful selection procedures with respect to L2 proficiency (3.4.4) and a classification of first languages in specific reference to the structures under investigation (see 3.5.3). The technique used to gain control over first language influence is unorthodox, yet promising. Indeed, in the past, quite a number of FFI studies have used a sample of mixed L1 backgrounds, mostly operating under the silent assumption that the L1 did not affect the results of their studies. However, given that transfer has frequently been demonstrated, and in the light of the assumption that L1 transfer is structure-specific rather than language specific (Odlin, 2003), this technique may on the one hand be an important requirement for dealing with mixed L1 backgrounds, and on the other hand go a long way in controlling for transfer effects.

The dependent measures that were used to assess progress were motivated by the constructs of explicit and implicit knowledge (see 3.3). An important reason to choose the untimed grammaticality judgement test to assess explicit knowledge was that it is the most widely used test in FFI research to measure gain. An additional advantage was that it calls upon explicit knowledge quite unobtrusively, thus minimizing potential retest effects. Because the participants were selected on the basis of not having knowledge of the target structures, they felt the test was quite difficult and for this reason the number of test items had to be kept small. Nevertheless, the test was found to be quite reliable (see 3.3.2, Table 3.6). Also, earlier findings were replicated, which is an indication of the test's validity. There is certainly room for improvement, though, and Ellis (2004)

provides a number of recommendations to enhance the validity of explicit knowledge tests. Most importantly, in addition to identifying the ungrammatical feature in the stimulus sentence, Ellis recommends to ask students to indicate how certain they are of their judgement, and whether they made their judgement on the basis of knowledge or intuition.

The most important purpose of the implicit knowledge test was that it obtained spontaneous written text from the participants without focusing them on the target structures. It can safely be concluded that the implicit elicitation technique that was used in this study was successful. In the beginning, the participants were quite concerned with providing correct answers, but once it was made clear that any answer was correct as long as they focused on the correctness of their formulations, most students actually enjoyed responding to the situations they were presented with. Even students that seemed to have adopted the strategy to give short answers (less scope for errors) were still found to use the target structures in their answers. Because of the elicitation techniques, data processing was fairly cost-efficient, in that a relatively small amount of data needed to be gathered as compared to sampling in real-life spontaneous settings. However, this testing method does require extensive piloting and fine-tuning in order to be certain that it elicits appropriately. Also, it should be noted that as soon as such kinds of elicitation techniques are used, particular contexts of use may be emphasized that are perhaps not as frequent in real time use. Reliabilities were calculated based on whether a correct form of the target structure was provided for each of the defined situations, and these reliabilities were found to be good (see 3.3.3, Table 3.7).

Because sampling and data processing were fairly cost-efficient, group sizes could be quite large. As a result, the statistical power to detect significance was considerable: effect sizes of f = .15 and higher could be detected. Sufficient statistical power is important in order to exclude the possibility that absence of significant effects – which were hypothesized – is due to an insufficient number of participants. Only for the analyses with regard to developmental readiness, the obtained statistical power turned out to be too small.

In short, this study has controlled for a number of features that may have introduced variation in earlier classroom FFI studies. Most importantly, the instruction was equal for all participants, and the exposure to the target structures during the instruction was highly regulated. Because of the random assignment to the experimental conditions, effects of potentially interfering variables were most likely mitigating rather than distorting.

5.3 Evaluation of theory

5.3.1 Introduction

In this section, the outcomes of this study will be evaluated in relation to the interface debate (2.2), the ideas on SLA expounded in 2.3, and earlier FFI research findings (2.4). This discussion will typically first recall into memory the theoretical framework and earlier findings, followed by a short summary of this study's findings, and a discussion of the meaning of these findings. This discussion is roughly structured according to the research questions. Thus, 5.3.2 discusses the findings in relation to the exploration of how L2 learners started to make use of the two target structures. Next, in 5.3.3, the evidence in relation to the interface debate is discussed. 5.3.4 discusses the different sources of individual difference: how they interact with instruction and how they impact upon grammatical development. Finally, in 5.3.5, discusses the importance of structure complexity in relation to grammatical development and instruction.

5.3.2 The nature of implicit L2 knowledge

In Chapter 2, the constructivist views on implicit second language knowledge were presented. The main tenet of the constructivist view is that grammatical knowledge of the L2 emerges in the course of acquisition. The development of L2 grammatical knowledge starts with the internalization of exemplars: words and phrases that contain grammar structures are stored as wholes in memory. L2 learners may use these chunks when producing language, but not in a computational way. Rather, they are retrieved from memory as such. Through subconscious reflection on the internalized exemplars, a process referred to as syntacticalization, L2 learners ultimately develop an open class grammatical system much like construction grammar. Learners may observe regular patterns in the intake, and start using such structures accordingly. Such patterns may still be limited in scope, in that they are related to specific lexical contexts. The constructivist view on the nature of L2 knowledge is closely related to the views on representation of L2 knowledge. Implicit knowledge is seen as represented in a complex and dynamic associative network, and implicit learning is essentially the frequency-based establishment of associations and the subsequent abstraction of grammatical categories. These do not come into existence overnight, but result from frequent exposure and a subconscious process of figuring language out. In

short, implicit learning is the inevitable and uncontrollable result of information processing (N. Ellis, 2002; Hulstijn, 2002).

One goal of this study was to investigate how L2 learners start to make use of developing L2 grammar structures. To this end, the use of the two target structures, the degrees of comparison and subordinate clauses, was investigated by means of the free written response data that was gathered over a period of three to four months. One aspect of use that was investigated was the notion of differentiated correct use across different types of realizations. Another aspect of use under investigation was the time it took L2 learners to start using the target structures in a flawless manner. The analyses were intended to be explorative and descriptive, and no hypotheses were formulated.

Differentiation in correctly using the target structure across different realizations was clearly demonstrated for SubCs. For this structure, substantial differences were found in correct use of conditional and causal subordinate clauses, amounting to 36 percentage points at T0, and remaining as large at T2. Conditional SubCs were used correctly more often than causal SubCs, and they were also used *before* causal SubCs. In addition, there were differences in correct use related to clause and verb phrase complexity: application of the structure in complex contexts meant higher rates of incorrect use. For the DoC, the findings were less clear. At T0, 80 percent of the forms were used correctly. This applied to all but one type of use: if comparatives were used in combination with comparative clauses, no more than 50 percent of the form was used correctly.

The findings with regard to SubCs suggest the application of a different word order in subordinate clauses to be lexically specific. Especially the difference in correct use between conditional and causal subordinate clauses was too substantial to be coincidental, and it clearly suggests that, in beginning stages of acquiring the ability to use SubCs, L2 learners relate the word order change that subordinate clauses require to the use of a particular subordinate conjunction. The question that rises is at what point – if at all – L2 learners do in fact connect the various types of subordinate clauses as belonging to one particular class of clauses with common formal characteristics. In other words, at what point is a more general grammar structure abstracted. The only relation that was found between the two types of SubCs studied, is that correct use of conditionals clearly preceded correct use of causals. If L2 learners had already recognized both types of SubCs as one class of structures, this difference should not have occurred. One can only assume that at one point, the connection is made. However, an account that sees the different types of SubCs as separately represented is tenable as well.

The absence of clear findings of lexical specificity in the use of the DoC may be due to a data analysis problem. Especially the differentiated use of comparatives in combination with comparative clauses is indicative of this. As pointed out in 3.3.3, L2 learners' intentions to express comparison may go unnoticed, because of a tendency to fail to mark adjectives for comparison while they do intend to express comparison. Because comparative clauses clearly betray the intention to compare, omissions can be – and were – counted as incorrect. However, for all other uses of the DoC such contextual clues are often not present, possibly resulting in unnoticed errors. For example, L2 learners writing "Deze radio is mooi" (This radio is nice) may actually have meant to write "Deze radio is mooier" (This radio is nicer). Because the L2 learners' true intentions could not be determined, such phrases were coded as 'no use'. Given the fact that the test items implicitly elicited comparison, and the fact that higher rates of elicitation were obtained in the pilot phase for more advanced students, it is likely that numerous instances of incorrect use were missed. As a result, the percentages of correct use of the DoC were overestimated, and conclusions about differentiation of use of the DoC in different contexts of appearance cannot safely be drawn.

For both target structures, progress was found to be gradual and slow. Progress was investigated by means of the stages of development proposed by White (1998). Four patterns of use were distinguished: no use, only incorrect use, variable (both correct and incorrect) use, and only correct use. Despite the average increase of correct forms in the free written response task, the participants were found to largely continue their ways of using the target structure: over 50 percent of the students demonstrated the same pattern of use from one testing moment to the other. If they did change their ways, they were found primarily to progress. For example, if students had at one point used correct forms of the target structures, they were hardly ever found to return to the use of only incorrect forms in subsequent tests. These results demonstrate that learning to use a grammar structure is not an instantaneous affair. There do not seem to be clear turning points that mark the acquisition of a particular structure. Rather, learning to use a structure correctly involves an extended period of time during which incorrect use and correct use co-occur, until at one point incorrect use disappears. Correct use seems to start in the more simple or salient contexts, and is gradually extended to more complex contexts.

This gradual learning process of replacing incorrect use for correct use agrees with the frequency-based language learning account as outlined in 2.3. Based on exposure to the target structures in the input, 'correct' representations of the

structure are established that gradually outweigh the 'incorrect' representations. It is important to understand that 'incorrect' here does not refer to faulty application of a structure, but to non-use. For the DoC, virtually all forms that were found to be and coded incorrect were examples of non-use: comparative markers were failed to use while contextual clues clearly indicated the intention to express comparison. Similarly, for SubCs, incorrect use hardly ever meant that L2 learners realized incorrect, non-existent sentence structure. Rather, sentence structure was left unaffected, and default main clause word order was used. The image that rises is that L2 learners have to replace a default structure that is being applied to express a broad set of meanings with specific structures associated with specific meanings. This process of replacement takes time, and during this process, L2 learners seem to have both correct and incorrect forms represented and available, alternating in a seemingly spurious way between them.

5.3.3 FFI and the development of explicit and implicit knowledge

The primary goal of this study was to investigate whether explicit instruction contributes to becoming proficient in using the second language grammar. Theoretically, this concern is addressed by the interface debate. Three different positions assign a different role to explicit knowledge in the course of developing implicit knowledge, the kind of knowledge that allows speakers to be proficient in the L2. In essence, the controversy in the interface debate evolves around the organisation of linguistic knowledge. All three positions make a distinction between explicit and implicit knowledge. For the strong interface position, knowledge of the L2 starts out as explicit and converts into implicit knowledge: one learning process is posited, covering several stages. Thus, there is a *direct* relationship between explicit and implicit knowledge. The weak interface position and the no interface position both argue explicit and implicit knowledge to be completely separate. The weak interface position still allows for the conversion of explicit knowledge into implicit knowledge, but argues that this depends on the kind of grammar structure. It also allows for delayed effects. The no interface position claims the two knowledge systems to be completely separate, and no conversion is possible. Potential effects of explicit knowledge are *indirect*, in that it may affect implicit learning processes.

In Chapter 2, this theoretical problem was approached from two different perspectives. On the one hand, recent SLA developments were evaluated in order to arrive at a better understanding of the constructs of explicit and implicit

knowledge. On the other hand, FFI research was scrutinized in search of evidence that explicit knowledge promotes the development of implicit knowledge. The outcome of this undertaking was that neither of the two approaches provides arguments in favour of a weak or strong interface. Based on this, it was hypothesized that there is indeed no interface between explicit and implicit knowledge: explicit instruction should not be more effective than implicit instruction in promoting the ability to use the language in spontaneous language use. Nor should it be less effective, because concomitant implicit learning effects will arise during the explicit instruction. These hypotheses were tested by comparing how explicit and implicit instruction affected the development of both explicit and implicit knowledge.

FFI and explicit knowledge

The results of this study replicate the predominant finding in FFI research that explicit instruction is more effective in promoting L2 knowledge as measured by explicit knowledge tests. For both target structures, significant immediate (i.e., from T0 to T1) explicit progress was observed when the rules of grammar were explicitly taught. Not much additional progress was observed once the instruction stopped. Implicit instruction was found to be effective only for one of the two target structures, subordinate clauses, from T0 to T2. Similar results have been reported in other studies using measures of progress that tap explicit knowledge. In such studies, implicit instruction is hardly ever found to be superior to explicit instruction: there are either no differences between explicit and implicit instruction, or explicit instruction works better. This is true, for example, for the studies reviewed in 2.4.5, but also Norris and Ortega (2000) – though not differentiating for the kinds of tests used – reported substantially higher effect sizes for explicit types of instruction over implicit types of instruction in their meta-analysis.

The results of this study suggest that there is an interaction between the kind of grammar structure taught and the kind of instruction received. For both target structures, explicit instruction was found to promote explicit knowledge quite effectively. Implicit instruction proved effect only for one of the two structures. For the simple morphological structure (the degrees of comparison), no effect of implicit instruction was observed. For the complex syntactic structure (subordination), on the other hand, implicit instruction ultimately (by T2) led to equal amounts of explicit knowledge. In previous studies, similar findings have been reported. DeKeyser and Robinson, for example, also found explicit

instruction to be more effective than implicit instruction for simple structures. In a similar vein, De Graaff found explicit instruction to work better for morphological structures, and found no differences between explicit and implicit instruction for syntactic structures. Based on this study and previous studies, then, the conclusion must be that EI is more effective in promoting explicit knowledge of simple structures. Implicit instruction promotes explicit knowledge only when the structure is complex.

Explicit instruction promoting explicit knowledge is according to expectation in the light of the definition of explicit knowledge as being declarative in nature and requiring deliberate effort to be learned. That implicit instruction may be just as effective is somewhat puzzling, and it is even more puzzling that this should hold only for complex grammar structures. Transfer of implicit knowledge probably explains why students in the implicit condition also improve their explicit knowledge: in 2.4.2, it was already pointed out that it is difficult to prevent students from using their implicit knowledge in explicit knowledge tests. Thus, the long term progress in explicit knowledge of the SubCs observed for students in the implicit instruction condition may reflect growth in implicit knowledge which students may revert to in absence of explicit knowledge. However, why this should have occurred only for the complex structure remains difficult to answer.

The growth of explicit knowledge was found to be durable. At T2, performance was equal to performance at T1 for both target structures, which means that the knowledge was not lost nor had any further knowledge been acquired once the instruction had stopped. This finding is in line with the definition of explicit knowledge proposed by R. Ellis (2004; see 2.3.5). Conscious, declarative knowledge is learned deliberately, and further growth without deliberate learning would be unexpected. In fact, if not regularly attended to, explicit knowledge is probably lost. A number of FFI studies have indeed reported loss of explicit knowledge at delayed post testing.

Implicit knowledge

No differences in development of implicit knowledge were found between explicit instruction and implicit instruction, irrespective of the kind of grammar structure concerned. Overall, significant progress was obtained by all students in both experimental conditions. In terms of effects sizes, progress was moderate. It is difficult to ascertain whether the instruction gave rise to much implicit learning as opposed to students who did not receive any focused instruction. For the

degrees of comparison, the control group data were contaminated. For the SubCs, both experimental conditions demonstrated better performance than the control group, but because the control group did gain some progress between T1 and T2, differences between the control group and the experimental groups missed significance.

The finding that there was no difference between the two experimental conditions concerning implicit progress was according to expectation. Based on an analysis of SLA theory and FFI research, this study departed from the no interface hypothesis. Actually, the no interface position suggests that methods stimulating implicit learning should outperform methods stimulating explicit learning: if the two knowledge systems are separate, then instruction targeting explicit knowledge cannot readily be expected to promote implicit knowledge. Nevertheless, the expected outcome was that there would be no differences between the two experimental conditions concerning implicit knowledge growth. This expectation was based on Hulstijn's suggestion that explicit instruction leads to concomitant implicit learning. The results for both target structures confirm this notion, in that both experimental conditions demonstrated equal implicit progress.

The findings of this study are also in agreement with previous findings. In 2.4.4, FFI research was scrutinized in search of evidence for an interface between explicit and implicit knowledge. The outcome of this analysis was that in two studies an advantage was found for explicit instruction over implicit instruction (Muranoi and VanPatten & Sanz). However, in both studies, the explicit instruction involved quantatively more input in comparison to implicit instruction, and it was argued that the findings cannot safely be taken as evidence of explicit instruction being more effective in promoting implicit knowledge. In the other studies mentioned in 2.4.4, no differences were observed between explicit and implicit types of instruction. This study replicates these latter findings, and underscores the importance of keeping the amount of exposure to the target structures during the instruction equal.

An important implication of these findings is that implicit learning will occur as long as relevant information is processed, irrespective of *how* the information is processed. In this study, care was taken to keep the exposure to the target structures in both experimental groups equal: even the sentences in which the structures appeared were literally the same across the conditions, although they were used either to explain and practise the target structure (EI), or to do text comprehension exercises (II) (see also 3.2). Despite the fact that the structures were processed for completely different purposes, there were no differences in

implicit learning results. This not only agrees with the notion of concomitant implicit learning during explicit instruction, it also confirms the idea of information processing being an uncontrollable process (see 2.3), which inevitably leads to learning. Without wanting to devaluate the importance of quality (i.e.: how the input is processed), quantity of exposure seems to be a vital aspect of learning.

The growth of implicit knowledge was found to be durable. The progress that was obtained between T0 and T1 was not lost between T1 and T2. However, a difference was observed between the two target structures. No further implicit knowledge of the DoC was obtained between T1 and T2, while implicit knowledge of SubCs continued to grow once the instruction stopped. Possible reasons for this difference will be dealt with in 5.3.5, where structure complexity is discussed.

5.3.4 Individual differences and the effectiveness of instruction

Another goal of this study has been to relate explicit and implicit learning to particular individual/personal qualities. On the one hand, such an investigation could uncover dependencies between the ability to use explicit knowledge to improve L2 performance and particular ID variables. On the other hand, such an investigation could provide further evidence for the separateness of the explicit and implicit learning and knowledge systems if ID's are found to be related differently to explicit and implicit grammatical development. On a theoretical level, ID's are quite commonly held to moderate the effectiveness of instruction. One of Krashen's claims when he put forward his no interface position was that only some learners are able to use their explicit knowledge as a monitor. In addition, he claimed that explicit and implicit learning are referenced by two different individual factors: attitude and aptitude. Similarly, Ellis argued that the effectiveness of instruction interacts with the L2 learners' stage of development. But also age and L1 background have been related to effective FFI (see Norris & Ortega, 2000, p. 422). Nevertheless, FFI studies relating the effectiveness of instruction to individual characteristics have been rather scarce, and basing predictions on earlier research is difficult. In this study, a number of potentially moderating learner variables have been included (developmental readiness, L1 similarity, attitude, aptitude, age, and cognitive style), with the objective to explore the relation between these variables and explicit and implicit progress.

Developmental readiness

Developmental readiness refers to the notion that L2 development is to some extent subject to fixed orders of acquisition. The implications for instruction are that instruction can only be effective when L2 learners are sensitive (i.e., developmentally ready) to the instruction because it matches their stage of acquisition. This study provides some evidence that FFI is more effective when L2 learners are developmentally ready. More precisely, developmental readiness (DR) was not found to be related to SubC progress, but it was found to be related to both explicit and implicit progress in the DoC. There were no interactions between DR and the type of instruction received. The meaning of these findings is that instruction in the DoC was more effective once the L2 learner is developmentally ready; how the instruction is realized – explicit or implicit – is inconsequential. It should be pointed out that for some of the analyses (those pertaining to the DoC and conditional SubCs) insufficient statistical power was reached, meaning that if significance failed to occur, this may have been caused by an insufficient number of subjects.

Two aspects about these findings are difficult to interpret. The first unexpected finding is that DR was found to influence the development of the DoC only. One may argue, though, that this is due to the differences between the target structures, and these will be discussed in detail in the following section. The second puzzling finding is that both explicit and implicit DoC gain were affected by DR. Given the definition of explicit knowledge (see 2.3.5), there is no limit to what can be learned explicitly at any given time, and explicit learning should not be constrained by DR. Nevertheless, the participants in this study that were developmentally ready improved their explicit knowledge more than students who were not yet ready.

Practical rather than theoretical grounds motivated the operationalization of DR. DR was defined in terms of emergence of correct use. The advantages are that emergence of correct use is not structure-specific and that it is clearly observable in the language of learners, even in classroom settings. In 3.5.2, it was already pointed out, though, that this operationalization is crude in the light of the underlying theory: knowledge of grammar structures develop according to fixed stages, and instruction should match these stages. This suggests that as learners move through the stages, instruction should be continuously adapted to match each stage. Such instruction is practically unattainable, which is why emergence of correct use was chosen. From a theoretical perspective, however, this is quite arbitrary, and because no relation was found between the development of implicit

SubC knowledge and DR, emergence of correct use does not seem to be a valid operationalization for all structures.

Another point needs to be made concerning the operationalization of DR. In this study, developmental readiness expressed 'advancedness' of the L2 learners' grammatical system in that two groups were created according to whether or not correct forms had emerged in the participants' production (see 3.5.2). The preliminary analyses have shown that the two groups that were thus created also differ significantly on a number of other variables: most notably grammatical accuracy and the C-test (see 4.4.2, Table 4.23). This is perfectly logical, as grammatical accuracy and the C-test can be considered expressions of 'advancedness of the linguistic system too, though not related to specific target structures. The question is what this means? Either, DR is simply a structure-specific indicator of proficiency that signifies at what point the input should be intensified, or DR groups the more talented and the less talented language learners together. This too would explain why the 'ready learners' obtained more progress. In addition, it explains the puzzling relation between DR and explicit grammatical progress.

This study does not provide evidence that the ability to use explicit knowledge to monitor implicit knowledge depends on developmental readiness. Perhaps instruction needs to be tuned more precisely to stages in the acquisition of grammar structures. The study does provide some evidence that instruction is more effective for those that have already demonstrated beginning correct use, and thus confirms findings by Williams and Evans (1998), who found that the greatest gains were obtained by those demonstrating knowledge of the target structures in the pre-test. The precise involvement of DR is difficult to determine. Either the construct successfully signifies a point at which instruction would be beneficial, or it merely groups the more talented and the less talented together.

L1 similarity

The only evidence of a positive effect of explicit knowledge on implicit grammatical performance was found in interaction with first language similarity, and only for the DoC. This finding will be further discussed in 5.3.6. None of the first languages were actually similar to Dutch, as there were no subjects with Germanic first languages present in the sample. Similarity in this study meant that the first language realizes the meaning expressed by the structures in focus in a similar way (see 3.5.3). The question is why L1 influence only observed for the DoC? The cause might lie in the differences between the two structures (to be

discussed in 5.3.5), but it more likely lies in the uniqueness of the inversion rule in Dutch subordination.

Most, if not all, languages have means to express the degrees of comparison. In Dutch, they are realized morphologically and sometimes periphrastically. Other languages similarly either use morphological marking, periphrasis or both, which makes comparison straightforward. In contrast, the inversion rule required in Dutch subordinate clauses is quite unique. Some other languages, such as German, have it too, but in many languages inversion in subordinate clauses simply does not exist. Only Armenian was found to have something similar, but there were only three speakers in the sample. Consequently, there was nothing similar in the participants' first language; and therefore, nothing to help understand the instruction. This may explain why no effects were found for the SubCs. Because subordination affects word order, default L1 word order was coded instead of inversion, but it was not found to affect the results.

Aptitude and attitude

The effects of a number of individual factors were studied in relation to explicit and implicit gain. To this end, explicit and implicit gain scores for both instruction conditions were related to the variables memory, grammatical sensitivity (GS), cognitive style, and motivation by means of correlation analyses. Research about the involvement of aptitude and attitude has often been triggered by Krashen's argument that attitude affects implicit learning, while aptitude affects explicit learning. Robinson and others have already pointed out that aptitude as it was measured by most aptitude test batteries was operationalized in a restricted sense, excluding components of aptitude that might relate to implicit learning. In this study, Krashen's claims have also been addressed. For aptitude, memory and GS were used, which both should correlate with explicit progress only. For attitude, motivation and cognitive style were used, as these two measures surfaced from a factor analysis on the teacher questionnaire.

Hardly any evidence was provided that these individual factors interact with instruction: if significant correlations were observed, they were quite low. In addition, the distances between the correlations observed for either instruction condition were never very large; they would fall within each others confidence intervals, which is why one cannot maintain that particular factors were differently involved in each type of instruction. However, some of the individual factors were found to be specifically related to either explicit or implicit DoC or SubC gain.

The absence of an interaction between instruction and aptitude confirms some of De Graaff's findings. In his study, he found that aptitude predicted progress in both instruction conditions equally. However, this study suggests a more complex relation between aptitude and progress. The findings suggest that the workings of aptitude are structure-specific: memory was related to correct use of the DoC in the free written production task, but not to correct use of SubCs. Conversely, Memory and GS were related to SubC grammaticality judgement performance, but not to DoC grammaticality judgement performance (see 4.4.4, Table 4.26 and 4.27). Such differences are unexpected, and can only be explained in relation to the nature of the target structures. This will be taken up in the following subsection (5.3.5).

On a theoretical note, the expectation that aptitude (as measured in this study) is positively correlated to performance in explicit instruction conditions is based on the notion that information is processed differently in the two conditions (see 2.3). If explicit and implicit instruction call upon different learning mechanisms, then that would constitute further evidence for the separateness of the two systems. The question that rises, though, is to what extent differences in processing during testing have been assessed, rather than differences in processing during the instruction. A grammaticality judgement task is bound to call upon mechanisms involving explicit handling of linguistic information, at least more than spontaneous production would. The relation between types of testing and aptitude has been examined in this study by relating memory and GS to five proficiency measures that were used as control variables. Both memory and GS were found to be significantly related to those measures that tested language proficiency in isolated contexts (the Cito ISK test and the C-test). However, no correlations were found between the aptitude measures and those proficiency measures based on spontaneous production (fluency and grammatical accuracy). It indicates that memory and GS indeed assess ed the involvement of explicit processing mechanisms. However, it is difficult to tease apart at what point these processes were involved: during the instruction or during testing.

With respect to motivation and cognitive style, no correlations were observed with either explicit or implicit progress. This study provides no evidence for positive affects of motivation or cognitive style on implicit learning. Given these results, the conclusion has to be that either the role of motivation and cognitive are limited, or their effects were too small to be captured, which is not unlikely in the light of the power limitations of the correlation analyses. Another possibility is that the teacher judgements that were used to assess motivation and cognitive

style were not adequately reflections of the learners' true motivation and cognitive style. The results, then, do not support Krashen's claim with regard to the role of attitude in implicit learning. However, this would have been unlikely in the light of the definitions of explicit and implicit knowledge. In fact, one can argue exactly the opposite to Krashen's claim; implicit learning has been characterized as an inescapable, uncontrollable process, and for this reason it seems unlikely that implicit learning would be restricted by aspects of attitude. Explicit learning, on the other hand, requires deliberation and conscious effort, and low motivation is likely to hinder such efforts. This study does not provide support for these claims either, but this may be due to insufficient statistical power or inaccurate assessment of attitude.

Age

Finally, the effect of age was assessed. Age was found to be related to explicit progress, in that the younger participants tended to obtain more explicit progress than the somewhat older learners. After a period of intensified input – either explicit or implicit, younger learners demonstrated more progress on the grammaticality judgement task than older learners. Two questions come to the fore: why is only explicit learning affected by age, and what explains the advantage of younger learners. Actually, a tremendous amount of literature has been devoted to the latter question (e.g., DeKeyser, 2000; Doughty, 2003; Robinson, 1997). It has been suggested that there may be a fundamental difference between learning at a young age and learning at an older age due to changes related to maturation or to the disengagement of an inborn language learning facility. These results actually suggest that a fundamental difference applies only to explicit learning. Although these results need replication, the repercussions for the fundamental difference discussion can be quite substantial. It would be particularly interesting to evaluate on what kinds of tests claims about a fundamental difference are based.

5.3.5 Structure complexity and the effectiveness of FFI

Fundamental differences in how different types of grammar structures are learned would seem unlikely according to constructivist theories of second language acquisition. The emergence of grammar being a frequency-based abstraction process applies universally to all grammar structures. In this light, the suggestion that different modes of learning (memory-based versus rule-based) take care of different types of grammar structures must be incorrect. However,

complexity in some form probably does affect the success of instruction. In fact, numerous complexity-related features of grammar structures have been proposed to affect the success of instruction: formal and functional complexity; scope and reliability; whether or not structures are developmentally constrained. Mostly, there is some evidence to support each of these proposals, but scientific comparison is difficult because if structures are classified according to all these distinctions, each structure is unique. Nevertheless, it seems that, if progress is assessed with measures calling upon explicit knowledge, simpler and more reliable structures can be taught explicitly with more success than complex and unreliable structures (see 2.4.5). It should be pointed out that simple structures in these studies were mostly also morphological structures. The findings for the impact of FFI on implicit progress are similar; the tendency is that morphological and formulaic structures are taught with more success than syntactic structures (R. Ellis 2002; see 2.4.3).

In design, this study involved two identical, parallel experiments in which two contrasting grammar structures featured. The goal was to discover to how instruction affects the acquisition of these structures. The target structures were chosen on the basis of practical and theoretical considerations. The most important difference is that the DoC structure is morphological, while the SubC structure is syntactic. In this study, differences were observed for the two structures a number of times. And for some of these, it seems that the differences are due to the difference in nature of the target structures. The question is to what extent these differences can be attributed to differences in the nature of the two structures.

One of the findings was that for the SubCs, differentiated use was demonstrated. Percentages of correct use were substantially higher for conditional SubCs in comparison to causal SubCs, and the same differentiation was observed for clause complexity and verb phrase complexity. In 5.3.3, this difference was already discussed and related to the possibility that incorrect use of the DoC may go unnoticed. However, another or an additional reason may lie in structure complexity. In fact, these findings provide suggestions towards how to define complexity. The very fact that the use of the DoC was not affected by the context suggests the structure has no impact upon the utterance beyond the adjective it is used with (except when comparative clauses are used). The differentiated use of the SubCs, on the other hand, suggests that the structure can be subdivided into different realizations of use that have to be learned more or less individually. These realizations are likely to be defined by the number of verbs the verb phrase consists of, and the constituents realized. Complexity, then,

may be a direct function of the number of different realizations a structure can appear in. This definition of complexity agrees with the persistent finding that morphological structures are taught with more ease than syntactic structures. Longitudinal research studying the development of using structures in spontaneous situations would be needed to gain further insights.

Another finding that may be better understood in the light of the differences in structure complexity is the finding that developmental readiness (DR) was found to affect DoC progress only. In 5.3.4, where DR was discussed, doubts were already expressed as to the validity of the operationalization. The construct of DR is actually based upon the notion of a staged acquisition process, and it was already pointed out that instruction may need to be continuously adapted to match the learners' stage of acquisition. However, perhaps sensitivity to instruction is better understood in terms complexity, meaning that DR should be determined in specific relation the acquisition of specific realizations of a structure rather than the structure as a whole. Assuming there are only few different realizations of the DoC, emergence of correct use accurately signifies sensitivity to instruction. For subordination, emergence of correct use was simply too crude in its operationalization, as no differentiation was made for the different ways the structure can be realized.

5.3.6 The interface issue

The findings of this study agree best with the no interface hypothesis. According to both the weak and the strong interface positions one would expect explicit instruction to be more effective; the no interface position as posited in this study would not expect any diffences between the two types of instruction, because both equally lead to implicit learning effects. Most compelling is the finding that explicit instruction was indeed not found to promote the ability to use the target structures in free writing any more than implicit instruction. The analyses clearly demonstrated that the participants who received explicit instruction had more explicit knowledge, yet this supposed advantage according to the strong and weak interface positions did not translate into higher rates of implicit learning. Even when the subject sample was cleared of those students that had not developed any explicit knowledge on the basis of the explicit instruction they received – after all, students may differ in their ability to learn explicit knowledge – no significant differences were found between students with and without explicit knowledge. In the light of these findings, it seems unlikely that explicit knowledge converted into implicit knowledge. Thus, it seems

justified to conclude that the strong interface position is incorrect; both types of knowledge are separately represented, and they result from separate learning mechanisms. The weak interface position is similarly not supported in as far as it supports the conversion of explicit knowledge into implicit knowledge.

Another finding that may be construed as evidence for the absence of an interface between explicit and implicit knowledge is the finding that the development of explicit and implicit knowledge was differently affected by the two types of instruction. For the development of linguistic knowledge as measured by grammaticality judgements, explicit instruction was found to lead to more knowledge gains than implicit instruction. If linguistic knowledge was measured implicitly, however, no differences were observed between the two types of instruction. For both target structures, significant implicit progress was observed, but there was no advantage of one type of instruction over the other. If the grammaticality judgements and free written response task had assessed the same kind of linguistic knowledge, this difference would be unexpected. Both tasks must therefore be considered to be measures of different kinds of linguistic knowledge.

Both the weak interface position and the no interface position allow for indirect effects of explicit knowledge upon implicit knowledge, meaning that having explicit knowledge may affect implicit learning processes. Ellis (1994a) also sees explicit knowledge as a facilitator of implicit knowledge in that it may help the L2 learner to notice and to notice-the-gap. Similarly, Krashen (1982) admits to potential indirect positive effects of explicit knowledge, because it may sometimes be used as a monitor or it may make the input more comprehensible (see also 2.2.2). In fact, this study provides a little bit of evidence of such an indirect positive effect. For the degrees of comparison only, L1 similarity was found to interact with the kind of instruction received. Students who received explicit instruction in the degrees of comparison and who had first languages that also realize comparison morphologically were found to significantly outperform implicitly instructed students and students with different L1's. An explanation for this result may be that explicit instruction made the participants aware of the similarities between their first and second language, reducing the perceived complexity of the structure, and consequently enabling them to apply first language processing strategies to their written L2 performance.

The findings with regard to L1 similarity may also suggest that the possibility of indirect positive effects of explicit knowledge upon the development of implicit knowledge may depend on the quality of the explicit knowledge. Making students aware of this similarity, although this was not done intentionally in this study,

may have improved the quality of the L2 learners' explicit knowledge to such an extent, that successful monitoring became possible. In this study, explicit knowledge was taught by means of the computer, and although great care was taken to teach the target structures effectively, the explicit knowledge these participants had may not have been good enough. It is true that teaching via the computer does limit the ways in which rules can be practised, and it does lack the interaction with a teacher who may perceive inadequacies in the students' knowledge where the computer would miss these.

All in all, this study lends credence to a no interface position, but it should be acknowledged that this study was not water-proof. There is a chance that this study has failed to establish the presence of an interface because of the target structures that were chosen. Another distinct possibility is that different types of explicit instruction aimed at facilitating input processing are more effective. Also, the delay in effects that the weak interface position expects may not have been long enough in this study. And finally, factors that were not included in this study, or that may have been inappropriately operationalized, may have covered up the existence of an interface. Nevertheless, the limited influence of explicit knowledge on the development of proficiency is in agreement with SLA theory and previous FFI findings. Is should be pointed out that these no interface claims are domain specific: they apply to the domain of syntax only, and its use in spontaneous settings. In addition, it remains to be seen to what extent these results can be generalized to older learners.

5.4 Implications for second language pedagogy

The ultimate goal of second language teaching is proficiency. This report has argued that, when trying to become proficient in their second language, L2 learners most likely stand very little to gain from having explicit knowledge of the rules of grammar. The question that rises is what role should be attributed to explicit instruction in second language acquisition curricula, especially in the light of the fact that the no interface position has been a reason to plea in favour of dismissing explicit types of instruction. It is important to stress that these findings do no support such pleas. This does study has not demonstrated the superiority of explicit types of instruction, but it also has not demonstrated the superiority of implicit types of instruction. Explicit instruction can be a valuable pedagogical tool, because it triggers concomitant implicit learning effects.

First of all, some researchers say the value of explicit instruction does not lie in the establishment of explicit knowledge, but in that it can promote input processing mechanisms. As already pointed out in the introduction (Chapter 1), So-called 'focus-on-form' research has concerned itself with the provision of instruction intended to shortly shift the focus to form during communication, arguing that the moment at which errors occur in communication is the moment at which (explicit) instruction would be most effective (e.g., Doughty, 2001b; Doughty & Williams, eds., 1998). Similarly, Processing Instruction (PI) aims at facilitating input processing mechanisms by pointing out to learners that their default processing strategies put them on the wrong track. As a note on the side, it should be remarked that this study contradicts this latter claim. The findings with regard to L1 similarity, however, suggest that explaining when default processing strategies *would* be effective is a good approach to explicit instruction, which seems to clash with VanPatten's claims. Nevertheless, this study has not addressed such claims, and leaves open the possibility that explicit instruction is indeed superior to implicit instruction in such cases.

Explicit instruction can be a valuable added ingredient in second language learning programmes provided they target implicit learning. There are in fact a number of reasons to resort to more explicit types of instruction every now and then. First of all, lessons in which explicit information is provided may form a good pretext to offer intensified form-focused input. Some structures may actually be quite difficult to supply in communicative settings of language use, and more explicitly focused language practice may then successfully lead to intensified input provision. Another reason to incorporate explicit teaching in the curriculum is simply that it meets the desires of many learners. As Krashen puts it: "There is no denying that that there is a certain satisfaction, for some of us, in knowing a conscious rule." (Krashen, 1981: p. 113). And this statement does not only apply to second language learners, but also to their teachers. Thus, explicit instruction may be valuable simply to keep class interesting and varied for both teachers and students. One complaint that was put forward by users of *Zebra* (Alons, Bienfait, & et al., 1999; and see 3.4.2) – in which grammar is offered implicitly only in the first two years – is that the exclusive focus on communication actually becomes tedious. A form-focused exercise every-now-and-then would be a welcome change for those who work with this textbook nearly everyday for at least a year.

It should also be acknowledged that explicit knowledge is not necessarily worthless. It may be valuable in particular circumstances. This study investigated two grammatical phenomena, and perhaps explicit knowledge of other

grammatical structure is useful to have. Also, explicit knowledge may not aid the acquisition of implicit knowledge, but once L2 learners have implicit knowledge of a particular grammatical phenomenon, explicit instruction be useful to have for situations of language use that allow for preparation. Also, the focus in this study was exclusively on the domain of grammar. No conclusions should be drawn about the value of explicit knowledge about other aspects of language, such as pronunciation, vocabulary, and communication strategies.

5.5 Conclusion

This study has contributed to the area of FFI research in a number of ways. First, counter to current thinking, this study has come to the conclusion that the value of explicit knowledge is limited. In an attempt to demonstrate the presence of an interface between explicit and implicit knowledge, no evidence of such an interface has been provided. In the experiment that was conducted, one group of students received explicit instruction in two grammar structures while another group received implicit instruction. In addition, a no-treatment control group was included. Having explicit knowledge of the target structures has not proved to be an advantage to L2 learners in that no differences in implicit knowledge were observed between the implicit and explicit treatment groups. Because the strong interface position supposes unconditional conversion of explicit knowledge into implicit knowledge, it is falsified by these outcomes. Indirect effects of explicit knowledge upon implicit knowledge cannot be ruled out; in fact, this study has found that explicit instruction effectively promoted implicit knowledge of the simple morphological target structure, if the L2 had a similar structure. The parallelism between the first and second language either simplifies the explicit knowledge to such an extent that monitoring becomes a possibility, or it allows for successful transfer to occur. No further interactions between instruction and grammatical progress were observed.

This study also gives rise to reconsider earlier FFI findings. In part, this study was motivated by a bias in FFI research towards the use of decontextualized knowledge tests at the expense of tests assessing the ability to use grammatical knowledge in natural settings. The outcomes clearly suggest that measures assessing grammatical knowledge in isolated contexts cannot readily be taken as representative of the learner's implicit grammatical knowledge. FFI findings need to be reconsidered according to how progress was measured. Given that the use of measures of natural language have been rare, replication studies will be welcome.

For example, as this study has shown, but also pointed out by the DeKeyser (2003), there are surprisingly few studies that actually compared explicit and implicit types of instruction. Most FFI studies have laboured under the assumption that explicit types of FFI are necessary, and have moved on to investigate the effects of different kinds of instruction in interaction with numerous other variables. Such studies, too, need to be considered in the light of the kind of measures used.

Unusual to FFI studies, even to those that have used measures of spontaneous language use, was the qualitatively oriented exploration of the free written response data. This analysis has proved to be quite rewarding, though. Perhaps the most important outcome was the clear differentiation in the ability to correctly use conditional and causal subordinate clauses. The finding suggests that language learners do not see the parallelism between the two manifestations of the structure, and they seem to learn each manifestation independently. On the basis of these findings, and in as far as possible, the analyses for subordination have been conducted separately for each type. These findings have also provided suggestions as to how to define structure complexity. Complexity might be directly related to the number of different manifestations a structure can appear in. Although it is not entirely clear what would constitute a single manifestation, it seems fair to conclude that syntactic structures generally have more than morphological structures. Thus, this explanation would explain why instruction has primarily been found to be effective with morphological structures.

This study emphasizes that L2 research is in desperate need of longitudinal studies that investigate how L2 learners start to use their developing L2 knowledge. Not only would this inform theories of knowledge and representation, it is probably also an important precursor for the kind of study conducted here. If one wants to make claims about the effect of instruction on grammatical development, it is important to understand precisely how particular structures emerge in the L2 learner's language. With simple frequency counts and with points of measurement that were quite far apart, this study has been able to uncover interesting features of learner language. Much can probably be learned if learner language is more systematically examined.

This study underscores Doughty's claim that: "the case for explicit instruction has been overstated." (2003, p. 274). However, these results should not be taken to mean that second language learning can do without instruction, as the no interface position has often been interpreted. Although in this study the effects

of instruction were not spectacular, it has been very clearly demonstrated the instruction is an important catalyst of learning (Larsen-Freeman & Long, 1991; Long, 1983). The purpose of instruction should be to intensify the input. The focus should be on meaningful interaction, but some explicit FFI would probably do more good than harm. Grammar instruction should not aim for the establishment of explicit knowledge. Rather, for grammar instruction to be optimally effective, it should be geared towards the provision of structures in realistic and functional settings, and teaching the structure in isolated contexts should be avoided. After all, the goal of second language grammar teaching is not conscious knowledge of grammar, but proficiency.