The effectiveness of comprehensive corrective feedback in second language writing
van Beuningen, C.G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 6

Conclusion

6.1 Introduction

The role of corrective feedback (CF) in the process of second language acquisition (SLA) is a topic that is receiving a lot of research attention. Sheen (2010a) attributed this continuing research interest in CF to “the significance it carries for both SLA theory building and language pedagogy” (p. 177). The series of studies presented in this thesis were carried out with this dual perspective in mind. Framed within a cognitive account of SLA, they set out to investigate the potential of comprehensive written CF in fostering second language (L2) development, and opted to translate empirical insights into both theoretical and pedagogical implications.

The first two investigations reported on in this dissertation are two quantitative, tightly controlled classroom-based studies incorporating pre-test, treatment, and (delayed) post-test sessions (cf. Chapter 3, N = 66 and Chapter 4, N = 268). They explored the effects of CF and two control treatments on the accuracy development of secondary school pupils learning Dutch as their L2. The third study is a qualitative exploration into the efficacy of written error correction (cf. Chapter 5, N = 4). The writing performances of four L2 learners (who were selected from among the participants of the large-scale study presented in Chapter 4) were subjected to an in-depth examination to advance the understanding of how and when individual learners benefit from written CF.

In this concluding chapter, I will synthesize the main findings produced by the three empirical studies presented in Chapters 3, 4, and 5 (cf. section 6.2). In addition, I will discuss their theoretical and practical implications (cf. sections 6.3 and 6.4). Finally, I will discuss the limitations of the present work, and sketch out some open issues and directions for future research (cf. section 6.5).
6.2 Synthesis of main findings

The empirical work presented in this thesis aimed at contributing to the error correction debate (e.g. Ferris, 1999; 2004; Truscott, 1996; 1999; 2007) by addressing six major issues: (1) the value of comprehensive CF as an editing tool, (2) the value of comprehensive written CF for L2 development, (3) the differential effectiveness of direct and indirect error correction, (4) factors mediating CF efficacy, (5) the potential negative side-effects of written error correction, and (6) how and when individual learners (fail to) benefit from CF. This section will summarize and synthesize the insights that the three empirical studies (cf. Chapters 3, 4, and 5) offered with respect to these topics. Finally, I will discuss some findings from Chapter 5 that do not directly relate to the above mentioned issues, but which are nevertheless relevant, as they demonstrate the added value of in-depth accuracy measures.

6.2.1 The value of comprehensive CF as an editing tool

The first objective of the present empirical work was to investigate whether comprehensive or unfocused CF (i.e. correction of all existing errors) has the ability to help learners develop more effective revision and editing skills. As did earlier research (e.g. Ashwell, 2000; Fathman & Whalley, 1990; Ferris, 1997; Ferris & Roberts, 2001; Sachs & Polio, 2007), the studies presented in Chapters 3 and 4 proved that comprehensive CF is indeed a useful editing tool, by showing that CF helps learners to enhance the accuracy of an initial manuscript during revision. Importantly, findings showed that CF has an added value above revision as such. Whereas pupils who were asked to revise their texts on their own (i.e. without any available feedback) were able to arrive at more accurate revisions, their revised texts were still significantly less accurate than those of pupils who received CF. These findings show that CF is a useful tool that can help learners in becoming more successful writers, which is important from a learning-to-write perspective on L2 writing.

6.2.2 The value of comprehensive CF for L2 acquisition

However interesting from a learning-to-write point of view, the observation that CF leads to increased accuracy during revision is relatively insignificant from an SLA perspective. The crucial concern of SLA research is the long-term effect of pedagogical interventions such as CF, or L2 development. Evaluating the SLA potential of CF therefore necessarily involves
CONCLUSION

“a comparison between two independently written works” (Truscott & Hsu, 2008, p. 293), instead of contrasting an initial text and its revision.

Whereas recent research already revealed that learners receiving selective or focused CF (i.e. correction of errors within a specific category only) are able to use the targeted feature(s) more accurately in new pieces of writing (e.g. Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007; Sheen, 2010b), the studies presented in this thesis were the first to show that unfocused CF leads to learning. The investigations reported in Chapters 3 and 4 found that pupils who received comprehensive CF made fewer errors in new pieces of writing than learners whose errors had not been corrected. Moreover, this positive effect of CF proved to be durable; it was not only visible one week after pupils received the corrections (cf. Chapters 3 and 4) but also retained over a four week period (cf. Chapter 4). Besides that, in-depth exploration of learners’ accuracy development (cf. Chapter 5) showed that the improvement brought about by CF was not item-based. One of the case-study participants (i.e. Emre) omitted several articles in his initial text. The CF he received on those errors led him to realize all obligatory determiners in his post-test writing, not just the ones that were corrected before. Similarly, the CF Dinesh received on constituent omissions resulted in improvement beyond the accurate use of the specific corrected items; even though, obviously, he did not use the same sentences (i.e. items) he received feedback on in his post-test writing, Dinesh’s newly written texts did not contain any incomplete sentences.

6.2.3 The relative efficacy of direct and indirect CF

Apart from exploring the overall effectiveness of comprehensive CF, the present work aimed at investigating the relative value of different comprehensive CF methodologies for learners’ accuracy development. Feedback interventions have often been categorized as either direct (i.e. teacher indicates errors and target forms) or indirect CF (i.e. teacher only indicates errors, not the target forms) types. The main feature distinguishing these two feedback approaches is the level of learner involvement in the correction process. When receiving direct corrections, the only thing a learner is expected to do during revision, is incorporating the target forms as provided by the teacher. Indirect CF, on the other hand, requires more active engagement from the part of the student. Using indirect CF can be seen as a problem-solving activity in which learners need to correct their own errors. The teacher only indicates that an error has been made (e.g. by means of underlining, error
codes, etc). In the present studies indirect CF took the form of error coding (i.e. the provision of different codes representing different types of errors, see Appendix C).

The results from the small-scale study reported in Chapter 3, led to the cautious suggestion that direct CF might be more beneficial to accuracy development than indirect correction. This conclusion was based on the observation that, during revision, pupils who received direct feedback committed significantly fewer errors than their classmates receiving indirect CF. Moreover, the direct CF group outperformed pupils in the two control groups when writing a new text, while the indirect CF group did not. However, the post-test difference between the direct and indirect feedback groups only took the form of a trend at a p-value of .06.

In section 6.2.4, I will explain that it might not be feasible to come to a unidirectional conclusion on the relative merits of direct and indirect CF. The studies presented in Chapters 4 and 5 revealed that the efficacy of the two feedback methodologies is dependent on the type of error that is targeted.

6.2.4 Factors mediating CF efficacy

There are many factors that could be expected to mediate the efficacy of (a specific type of) CF. The benefits of error correction might be dependent, for example, on the nature of the targeted error, learners' attitudes towards the provided feedback, or the goal a teacher pursues by providing CF. The studies presented in this thesis explored the effects of three such potentially mediating factors: (1) error type, (2) learners' educational level, and (3) the topic of the task on which feedback is given.

Error type

It has been argued that different types of errors might not all be equally amenable to CF (e.g. Ferris, 1999; Truscott, 1996), because morphological, syntactic, and lexical errors represent gaps within different domains of linguistic knowledge (e.g. Schwartz, 1993). The studies presented in Chapters 4 and 5 therefore explored the effects of CF on errors within different domains. Whereas Chapter 4 contrasted grammatical and non-grammatical problems, Chapter 5 took a more in-depth look into the CF responsiveness of separate error types. Both studies furthermore investigated if the nature of a targeted error interacts with the efficacy of direct and indirect correction.
Grammatical and non-grammatical errors
Chapter 4 opted to test Truscott’s (2001; 2007) hypothesis that error correction might lead to improved accuracy of non-grammatical features, but that grammatical errors are insusceptible to CF. This study’s analyses therefore distinguished between grammar errors (e.g. morphosyntactic errors) and errors outside the grammatical domain (e.g. spelling errors). Results refuted Truscott’s claim by showing that CF enabled learners to improve both their grammatical and non-grammatical accuracy over time.

Separate error types
It could be expected that within the two broad domains distinguished in Chapter 4, separate error types still differ in their level of CF responsiveness (e.g. Ferris, 1999; Truscott, 2001). The multiple case-study presented in Chapter 5 therefore provided a more in-depth exploration of the amenability of different types of errors to written CF. The study showed that CF facilitated accuracy development on a broad range of linguistic features (e.g. determiner usage, agreement, sentence structure, inflection, and punctuation), both on a local level (i.e. within one constituent) and on a more global, sentential level. These findings led to the conclusion that syntactic, morphological, lexical, and orthographical errors all benefit from CF.

Results revealed one situation in which CF might be unable to lead to learning. One of the case-study participants did not show any sign of feedback retention when CF targeted one of the most frequent errors in his work (i.e. a common Dutch spelling problem). It was hypothesized that extremely persistent and potentially fossilized errors might be too deeply entrenched in a learner’s interlanguage system to be susceptible to CF.

Interactions between error type and CF methodology
Both Chapters 4 and 5 aimed at exploring if the relative efficacy of direct and indirect CF is dependent on the type of error that is targeted. Findings suggested that this is indeed the case. The study reported in Chapter 4 showed that only direct correction promoted grammatical accuracy development. Pupils’ non-grammatical accuracy, on the other hand, proved to benefit most from indirect CF; in the non-grammatical domain, the effect of indirect CF was more durable than that of direct CF.

Findings from the qualitative study presented in Chapter 5 brought to light that the effectiveness of the different CF methodologies and error type are even more intricately related. As Ferris (1999) predicted, the efficacy of indirect correction showed to be dependent on whether or not a targeted feature is rule-based in character. The case-studies
revealed that errors in the use of rule-based features – article omissions for example – are suitable candidates for self-correction based on indirect CF. However, indirect CF proved to be unsuccessful in remediying linguistic problems that are not rule-governed, such as lexical errors. It was argued that in those cases where a learner cannot depend on a clear rule when trying to deduce a target structure, indirect CF might not be explicit enough to be beneficial. The qualitative data also confirmed Ferris’ hypothesis that more explicit feedback types, such as direct CF, might be more effective in developing the accurate use of idiosyncratic features; one of the case-studies illustrated that direct CF did foster accuracy development of a non-rule-based feature.

At first sight, the findings from Chapters 4 and 5 seem to be contradictory with respect to the effect that direct and indirect CF have on lexical errors. Being categorized non-grammatical errors in Chapter 4, indirect correction was implicitly claimed to be most effective in addressing learners’ lexical problems. Conversely, Chapter 5 suggested that indirect CF provides learners with insufficient clues to self-edit their lexical errors because they are idiosyncratic in nature. A closer look at Chapter 4’s non-grammatical error category, and its amenability to direct and indirect CF, is needed in order to unravel this alleged disagreement in findings. When considering the distribution of errors across separate error types, it becomes clear that lexical problems only made up 13 percent of the total number of errors within Chapter 4’s broad non-grammatical error category. When subsequently removing the lexical errors from the measure of non-grammatical accuracy and rerunning the analyses on the delayed post-test data, another important observation can be made. The original delayed post-test analyses reported in Chapter 4 (cf. section 4.6.4) already showed that indirect correction was significantly more effective in remediying learners’ non-grammatical errors than writing practice or self-correction without CF. The difference between the direct and indirect treatments, however, just failed to reach significance (p = .06) in the initial analyses. Interestingly, this difference between the two CF methodologies (in favor of indirect CF) did become significant (p = .04) after the lexical errors were taken out of the non-grammatical error category. What the study in Chapter 4 thus showed, is that non-grammatical errors outside the lexical domain profit most from indirect correction. Hence, the new analyses were able to resolve the conflict in the findings from Chapters 4 and 5.

Educational level
The study reported in Chapter 4 furthermore investigated the influence of learners’ educational level on CF effectiveness. From a pedagogical perspective this is an interesting
issue, because for teachers it is useful to know if learners across different educational levels respond differently to (a specific type of) written CF.

Additionally, exploring the potentially mediating effect of learners' educational level on feedback efficacy might lead to theoretical implications. It has been claimed that factors such as learners’ levels of language proficiency and meta-linguistic awareness influence the degree to which pupils are able to benefit from error correction in general, and indirect CF in particular (e.g. Ferris, 2004; Hyland & Hyland, 2006). As explained in Chapter 2, pupils’ educational level was assumed to be indicative of their level of (meta-)linguistic competence.

Chapter 4 corroborated the expectation that the higher-level pupils outperformed the lower-level learners on the different linguistic measures used in the study (i.e. measures of written accuracy, lexical and structural complexity, and vocabulary knowledge). However, no significant interactions were found between the efficacy of the different CF treatments and learners’ level of educational. One possible interpretation of this finding is that it rebuts the hypothesis that CF is more beneficial to learners with higher levels of meta-linguistic awareness (e.g. Ferris, 2004; Hyland & Hyland, 2006; Sheen, 2007). However, as will be discussed in section 6.5.3, it might also be that the presumption that learners’ educational level is related to their level of (meta-)linguistic competence is not a valid one.

Task topic
Writing researchers have suggested that it is not just learners’ level of L2 proficiency which determines the quality of the texts they produce. Other factors, such as a task’s topic, also proved to contribute to students’ writing performance (e.g. Meuffels & Van den Bergh, 2005; Schoonen, 2005). In this line of reasoning, it might be conceivable that learners’ knowledge about, interest in, and/or familiarity with the topic of the text they receive feedback on, also influences their motivation or ability to benefit from the provided CF. Chapter 3 therefore explored this potential mediating effect of a task’s topic by including writing tasks on two different themes. Results showed that the different topics of the two tasks used in the present study did not influence the extent to which learners benefited from the CF they received on their writing.

---

1 It needs to be noted that, in this type of studies, linguistic accuracy only constitutes one aspect of writing performance (which is most commonly assessed by means of holistic writing scores).
6.2.5 Potential harmful side-effects of CF

CF opponents proposed that, besides being ineffective, error correction could be expected to have a detrimental effect on learners’ accuracy development (e.g. Krashen, 1982; Truscott, 1996). The first argument raised to support this claim, is that correction and revision are too time-consuming, and that valuable class time could be spent more effectively on extra writing practice. Secondly, it was argued that the single goal of CF is making learners aware of the errors they committed. This awareness was thought to lead learners to avoid making the same errors in future writing, and therefore to result in less complex language use. The present research explored these two potential harmful effects of written CF by (1) comparing the performance of pupils receiving CF to that of learners who were given an extra opportunity to practice their writing skills, and by (2) assessing the effect of CF on the complexity of learners’ writing.

The results presented in this thesis refuted both of the above mentioned claims. The findings in Chapters 3 and 4 showed that pupils who received CF outperformed learners who were allowed an extra opportunity to practice their writing skills. Moreover, the results reported in Chapter 4 proved that CF did not lead to avoidance; CF did not result in a reduction of the lexical and/or structural complexity of learners’ writing.

6.2.6 How and when individual learners (fail to) benefit from CF

Chapter 5 subjected the sequential writing performances of four L2 learners of Dutch to an in-depth accuracy analysis to be able to gain insights into how and when written CF affects individual L2 learners’ accuracy development. Even though only based on the performances of four pupils, and caution should be taken in generalizing the findings, this study was able to lead to two interesting observations.

First, findings suggested that the potential of CF in yielding a learning effect might be dependent on a learner’s level of CF uptake (where uptake was operationalized as the successful revision of an erroneous utterance based on CF). The case-studies showed that in the majority of cases where pupils committed errors in their post-test writing on which they already received feedback, this lack of CF retention could be traced back to a lack of CF uptake during revision.

Secondly, qualitative analyses of learners’ errors and corrections provided some clues on why pupils sometimes fail to take up the CF they are provided with. It was suggested in Chapter 5 that CF instances targeting errors which are low salient – either in terms of their physical size or because they do not greatly affect the comprehensibility of a text – might be relatively easy to overlook, and will therefore not always be taken up.
Moreover, the case-studies seemed to suggest that affective factors, such as the way in which learners perceive a revision task (e.g. Storch & Wigglesworth, 2010), might influence learners’ level of CF uptake.

### 6.2.7 The value of in-depth accuracy measures

It has been suggested that global measures of accuracy improvement (e.g. error rate comparison) might fail to give an adequate and complete picture of the effects of CF on L2 learners’ accuracy development (Bruton, 2009a; Storch, 2010). The findings reported in Chapter 5 proved that this is indeed the case, and that in-depth error analyses could be considered a valuable supplement to the more common quantitative method of assessing CF effectiveness. Besides providing additional insights into the correctability of different error types and learners’ engagement with CF (as was explained in sections 6.2.4 and 6.2.6), the detailed approach also led to the conclusion that the effectiveness of CF is underestimated by overall error rate comparison. This conclusion was based on the following observations. First, global scores were shown to partially mask learners’ progress by including errors in new texts that were not related in any way to the errors corrected in the initial texts. As Bruton (2009a) rightly argued, CF could not be expected to promote the accurate use of features it did not target. Secondly, in some cases, learners only showed retention for part of a feedback instance. CF targeting article omissions, for example, led one of the case-study participants (i.e. Emre) to realize the obligatory determiners in new pieces of writing, but not the correct ones (i.e. Dutch differentiates between neuter and non-neuter determiners). This type of partial acquisition of a corrected feature, however, is not represented in global accuracy scores because such measures adopt all-or-nothing criteria (i.e. accurate/inaccurate) (Bruton, 2007; 2010).

### 6.3 Theoretical implications

This section evaluates how the findings presented in this thesis add to the theoretical understanding of the role of written CF in SLA. It will successively discuss (1) the value of CF to accuracy development, (2) implications relating to the role of attention in CF processing, (3) the relation between CF uptake and interlanguage development, and (4) the level of CF explicitness that is required for SLA.
6.3.1 CF and accuracy development

As was explained in Chapter 2, one's conceptualization of (the relation between) implicit and explicit L2 knowledge and the value attributed to explicit knowledge as such, are reflected in one's expectations with respect to the ability of CF to aid SLA. Whereas implicit L2 knowledge could be defined as knowledge of the language, explicit knowledge constitutes knowledge about the L2 system.

Scholars opposing the usefulness of CF (e.g. Krashen, 1985; Truscott, 1996) have claimed that, at best, error correction may lead to the development of explicit knowledge, and that such knowledge about the language is of very little value in itself. The position that explicit knowledge will never become implicit, then leads those CF contestants to conclude that CF has no role in L2 development.

The more common view, however, is that both implicit and explicit L2 knowledge are important links in the SLA process (e.g. DeKeyser, 1998; Doughty & Williams, 1998; Hulstijn, 1995; Hulstijn & Schmidt, 1994; Long & Robinson, 1998; McLaughlin, 1990; Schmidt, 1990; Schmidt & Frota, 1986; Swain, 1985), and that CF can be beneficial to interlanguage development, either by assisting the proceduralization of explicit knowledge (e.g. DeKeyser, 1998) or by fostering cognitive processes such as noticing (the gap) and hypothesis testing (e.g. Swain, 1991).

Whereas it is still an empirical question whether learners store CF as implicit or explicit knowledge (Bitchener & Knoch, 2010a), the present research (together with other recent written CF studies) does counter the claim that CF can only lead to 'pseudolearning' (Truscott, 1996). The fact that learners still benefited from the CF they received when writing new texts, demonstrates that error correction has the potential to lead to accuracy development. Importantly, these CF invoked accuracy gains proved to be durable; Chapter 4 showed that the observed positive effects of written error correction were retained over a four-week period\(^2\), and therefore refutes Truscott's (1996) idea that CF can only lead to a transient form of L2 knowledge. What makes the present findings even more promising is that the studies reported on in this thesis used realistic, communicative writing tasks which invited learners to focus on content rather than on language as an object. As pointed out by several researchers (e.g. Long, 2007; Ellis, 2010), such free constructed responses arguably afford the most valid measure of language development.

\(^2\) Bitchener & Knoch (2010a) were even able to show that the beneficial effects of written CF were retained as long as ten months after the feedback had been provided.
6.3.2 CF and attention

Attention is a crucial notion within cognitive SLA accounts. It is generally believed that only the subset of the input that is attended to is available for further processing (e.g. Robinson, 2003; Schmidt, 1990; 1994; 2001; Sharwood Smith, 1993; VanPatten & Cadierno, 1993). One of the hypothesized essential properties of attentional resources is that they are limited (e.g. Schmidt, 2001). Two related implications that follow from the present work concern the attentional demands associated with comprehensive CF and learners’ allocation of attentional resources to written accuracy and complexity.

Attentional demands of comprehensive CF

Advocates of a focused CF approach have suggested that targeting a (couple of) specific error type(s) should be expected to be more effective than comprehensive or unfocused error correction (e.g. Bitchener, 2008; Ellis et al., 2008; Sheen, 2007). In their view, comprehensive CF might fail to facilitate SLA because L2 learners have a limited processing capacity. Those in favor of focused correction argue that asking learners to deal with CF targeting a broad range of linguistic features at the same time might lead to a cognitive overload, and thus prohibit feedback processing.

In showing that comprehensive CF has the potential to yield sustained accuracy improvement, the present findings do not support the hypothesis that processing unfocused CF is too cognitively demanding. Instead, they suggest that learners have enough attentional resources available to attend to many different linguistic features within one text. As proposed in Chapter 2, it might be the offline character of writing that prevents learners from becoming cognitively overloaded when presented with unfocused corrections (e.g. Sheen, 2010a). What is important to note, however, is that the texts pupils wrote in the present studies were rather short (i.e. around 120 words). It is an empirical question if comprehensive CF is also an effective means of addressing longer texts.

Attention, accuracy, and complexity

Limited capacity models of attention hypothesize that focusing learners’ attention on accuracy (i.e. by providing CF) will lead to a reduction of the linguistic complexity of learner output (e.g. Skehan, 1998; Skehan & Foster, 2001). Those models assume that accuracy and complexity draw upon a single pool of attention and are therefore in constant competition for attentional resources.

The present findings are not in line with the accuracy/complexity trade-off predicted by single-resource accounts of attention. Chapter 4 showed that learners were able to
improve the accuracy of their writing without sacrificing on the end of structural or lexical complexity. It could therefore be suggested that these outcomes support a multiple-resource perspective on attention (e.g. Robinson, 2003; 2005). Within such a model, accuracy and complexity are not presumed to be in competition because these two performance dimensions of linguistic form are thought to be closely connected.

An alternative interpretation of the findings is also possible. It is feasible that it was only during text revision that pupils’ attention was explicitly focused on accuracy, since that was the moment in which they were made aware of their non-target-like language use (i.e. by means of CF). During revision, all attentional resources could be allocated to accuracy because all other features (i.e. content, complexity) had already been taken care of. When writing new texts, on the other hand, pupils’ attention was not explicitly drawn to accuracy issues. The fact that the writing tasks used in the present studies were communicative in nature without any inherent focus on language form, might have allowed learners to allocate minimal attention to accuracy in the post-test sessions, thus leaving sufficient attentional capacity to spend on the linguistic complexity of their writing. In this line of reasoning, one would not expect any accuracy/complexity trade-off effects, because there was never a need to pay focal attention to both accuracy and complexity simultaneously.

Unrelated to attentional capacity issues, but connected to the interrelationship between accuracy and complexity, is Truscott’s (1996) claim that CF may lead to avoidance of complex language use. He stated that by inducing learner stress and anxiety of committing the same errors in future writing, CF could make learners avoid the erroneous constructions when writing a new text. By demonstrating that CF did not lead to simplified writing, the findings presented in Chapter 4 rebut Truscott’s claim.

6.3.3 CF uptake and SLA

Uptake can be defined as a learner’s response to a CF instance or, as Lyster and Ranta (1997) put it, uptake is “what the student attempts to do with the teacher’s feedback” (p. 49). Uptake is successful when the original error is repaired. Although successful uptake in itself does not constitute evidence of interlanguage development (e.g. Ellis et al., 2001; Long, 2007), several researchers have proposed that a learner’s level of CF uptake might be predictive of L2 acquisition (e.g. Ellis & Sheen, 2006; Lightbown, 1998; Loewen, 2004; Lyster, 1998; Sheen, 2004). In their view, the fact that a learner was able to come to a

---

3 It has to be noted that the study reported on in Chapter 4 only included one measure of lexical complexity (i.e. Guiraud’s index) and one measure of structural complexity (i.e. subordination index).
CONCLUSION

target-like reformulation based on CF gives reason to believe that noticing has taken place, which, in turn, facilitates SLA (e.g. Schmidt, 1990; 2001).

The qualitative investigation presented in Chapter 5 suggests that – at least for the four learners in this study – CF uptake was indeed a good predictor of the success in acquiring a target form. As observed in both Chapter 4 and Chapter 5, in general, CF was able to trigger sustained gains in the accurate use of a targeted feature. When CF failed to lead to accuracy development, however, this lack of feedback retention could be traced back to a lack of successful CF uptake during revision. Interestingly, the observations reported in Chapter 5 also demonstrated that CF uptake forms no guaranty for long-term acquisition. It was shown that CF targeting an exceptionally deep-rooted and potentially fossilized error might not result in retention, even when the correction is taken up during revision. This finding thus validates the claim that the level of successful uptake, although predictive of acquisition, was not a reliable developmental measure in itself (e.g. Ellis, Basturkmen, & Loewen, 2001; Long, 2007).

6.3.4 The required level of CF explicitness

As was explained in Chapter 2, researchers have advanced theoretical arguments in favor of both direct and indirect CF methodologies. For example, indirect correction has been claimed to have a greater potential to foster SLA than direct CF because indirect CF engages learners in a more profound form of language processing (e.g. Ferris, 1995; Lalande, 1982). On the other hand, direct CF could be argued to be superior to indirect feedback forms because it presents learners with the kind of explicit information that is needed for cognitive learning processes, such as noticing and hypothesis testing (e.g. Bitchener & Knoch, 2010b). The studies discussed in Chapters 4 and 5 of this dissertation provide support for both of these hypotheses, and suggest that the required level of CF explicitness is dependent on the type of error targeted.

Chapter 4 showed that the beneficial effects of the less explicit, cognitively demanding indirect approach are more durable than those of direct correction when CF targets linguistic features that pupils are able and confident to self-correct, that is their non-grammatical errors (e.g. spelling errors, punctuation errors). However, this study revealed that feedback needs to be more explicit (i.e. direct) in order to lead to the development of grammatical (i.e. morphosyntactic) accuracy.

The qualitative study presented in Chapter 5 furthermore revealed that the needed level of CF explicitness depends on the idiosyncrasy of the targeted error. Whereas rule-based errors (e.g. inflection errors) showed to be suitable candidates for self-correction
based on indirect CF, direct CF proved to be more beneficial to learners’ errors in features that are not rule-governed (e.g. lexical errors). This finding is in line with DeKeyser’s (2003) suggestion that arbitrary form-function mappings will require more explicit learning processes. In those cases where a learner cannot depend on a clear rule when trying to deduce a target structure indirect CF might not be explicit enough to be beneficial. Therefore, maximally explicit feedback methodologies – such as direct CF – could be expected to be most effective in developing the accurate use of idiosyncratic features.

The observation that different error types show differential levels of responsiveness to direct and indirect CF methodologies, is in line with the claim that no single form of CF can be effective in addressing all linguistic error types (e.g. Ferris, 1999; Truscott, 1996), because morphological, syntactic, and lexical errors represent gaps within different domains of linguistic knowledge (e.g. Schwartz, 1993). It is important to note, however, that all error types have proven to be correctable, and that Truscott’s (2001; 2007) hypothesis that grammatical errors are insusceptible to CF could thus be rejected.

6.4 Pedagogical implications

Even though it is important to keep in mind that it is unsure if the findings of the present experiments would hold in a real-world class situation (cf. section 6.5), it is still worth considering some pedagogical implications from the studies presented in this thesis. This section discusses (1) the value of CF for language learning, (2) the efficacy of authentic, comprehensive correction, (3) the importance of revision activities, (4) the complementary value of different CF methodologies, and (5) the potential of CF within content-focused contexts.

6.4.1 The value of written CF for language learning

The main and foremost suggestion that follows from the present findings (as well as from other recent written CF studies) is that CF is a useful instrument that L2 teachers can employ to help learners improve their written accuracy. CF has shown to be valuable both as an editing tool and as a SLA intervention; it does not only enable students to improve a particular piece of writing during revision, but learners still benefit from CF when writing a new text. Just one feedback treatment already proved to be effective in constituting long-term accuracy gains. These beneficial effects could only be expected to be greater if learners are offered CF on additional occasions.
6.4.2 The efficacy and ecological validity of comprehensive CF

A second important pedagogical contribution of the empirical studies reported in this thesis, is that they revealed that comprehensive written CF has the potential to yield a learning effect. Until now, the only robust evidence on the long-term effects of written CF came from studies exploring the efficacy of focused correction, that is CF targeting one specific type of errors (e.g. Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009; Bitchener & Knoch, 2010a; Bitchener & Knoch, 2010b; Ellis et al., 2008; Sheen, 2007; Sheen, 2010b). However important these findings may be for SLA theory, some scholars have questioned their practical applicability (e.g. Ferris, 2010; Storch, 2010). They pointed out that, when providing CF, teachers usually opt to improve the overall accuracy of their students’ writing, not just the use of one specific linguistic feature. This might be particularly true in contexts where the main pedagogical focus is on communicating content rather than on language as an object (e.g. Anderson, 2010). It has therefore been claimed that comprehensive CF is a more authentic feedback methodology (e.g. Anderson, 2010; Ferris, 2010; Hartshorn et al., 2010; Storch, 2010). By exploring the efficacy of comprehensive CF, the present studies – even though still experimental in set-up – tried to reflect real classroom conditions in terms of the type of feedback provided. The finding that an authentic CF methodology such as comprehensive error correction leads to L2 acquisition could thus be considered of great relevance to teachers.

6.4.3 The importance of revising opportunity

A third implication that can be drawn from the work at hand is that teachers should allow learners the opportunity to revise their texts based on the provided feedback. Asking students to perform revision activities could be expected to foster SLA because “producing the correct form may help learners automatize their L2 production” (Loewen, 2004, p. 157). Moreover, producing accurate revisions could be considered a manifestation of pushed output, and should therefore be expected to promote L2 acquisition by triggering noticing (the gap) and hypothesis testing (Swain, 1985; 2005). Chapter 5 provided evidence of these potential advantages of text revision by showing that CF instances were more likely to lead to improved accuracy in new texts when learners were able to use them effectively during revision (i.e. when they showed signs of CF uptake).

In relation to the above, it could be considered important to make students aware of the goals and value of CF provision and revision activities, and to stimulate learners’ motivation to engage with the provided feedback. One of the case-studies reported in Chapter 5 showed, for example, that the potential beneficial roles of CF-based revision are
not served when learners misperceive its goals. Other researchers have also attested that affective factors – such as attitudes towards (a certain type of) CF, perceived goals, motivation – may influence the success of CF interventions (e.g. Bruton, 2009b; Goldstein, 2006; Hyland, 1998; Storch, 2010; Storch & Wigglesworth, 2010; Swain & Lapkin, 2003).

6.4.4 The complementary value of different CF methodologies
As was discussed in sections 6.2.4 and 6.3.4, the extent to which learners are able to benefit from direct and indirect CF methodologies, might be dependent on the type of error targeted. It was argued that explicit forms of CF, such as direct correction, are more effective in solving complex (e.g. morphosyntactic) or idiosyncratic (e.g. lexical) errors, while indirect CF might lead to more durable accuracy improvement when targeting relatively simple features (e.g. punctuation). Moreover, Ferris (2010) proposed that direct and indirect feedback methodologies might serve different goals. When the focus is on improving learners’ L2 composition skills, and CF is intended to help learners in becoming self-employed writers, indirect CF methods might be considered more useful because they demand a more active form of learner engagement. When mainly opting for language learning, on the other hand, direct correction might be judged the most beneficial approach because it provides the kind of efficient and explicit input necessary for L2 acquisition. I conclude that, depending on teachers’ goals and the types of errors that need to be attended to, direct and indirect CF interventions can complement each other.

The same holds for focused (or selective) and unfocused (or comprehensive) correction strategies. Since both CF types have now been shown to be effective, teachers can choose which approach they think is most appropriate in a specific situation. It might be useful, for example, to alternate comprehensive CF methodologies with intensive, focused correction when targeting linguistic features that have proven to be persistently problematic for a given learner.

6.4.5 The CF potential within a content-focused context
The present empirical work has shown that providing written CF is an effective means of drawing learners’ attention to language form in a content-focused context (i.e. biology class). One of the implications that follows is that time spent on correction of written products could be considered well-invested, even when learners’ L2 development is not the primary concern of the context the CF is provided in.

Albeit only based on anecdotic evidence, I would claim that this finding is important because content teachers might often be reluctant to devote time and attention to language
related issues; even in the context of the present research, where schools' policy prescribed a language sensitive approach to content teaching (e.g. Van Eerde & Hajer, 2008), many teachers felt that providing language-related feedback was the task and responsibility of the Dutch language teacher⁴. Additionally, some of them failed to see how the provision of CF would benefit their teaching, and stated that the language sensitive character of their pedagogical approach mainly involved extra glossing of content-specific vocabulary⁵.

However, research has shown that the language problems L2 learners have to deal with, go deeper than lacking the appropriate content-related vocabulary (e.g. Prenger, 2005; Van den Boer, 2003), and signaled a need for attention to linguistic form and accuracy in content-focused contexts (Schooten & Emmelot, 2004). The studies reported in this dissertation showed that written CF can be one of the tools content-teachers can use to fulfill this need.

One more argument in favor of providing CF within content-focused contexts and not just in language classrooms (within the same educational program), resides in the notion of transfer-appropriate learning (Segalowitz, 1997; 2000). It could be expected that feedback provided in a content-focused class is more beneficial to pupils’ future accuracy performance in this same context than CF provided in a fully language oriented lesson; what is learned from processing feedback in a certain context is conceivably more transfer-appropriate than the type of knowledge gained from CF provided in another setting.

### 6.5 Limitations and directions for further research

Although the studies presented in this thesis provide clear evidence in favor of written comprehensive CF, there are a number of limitations to the present empirical work that need to be acknowledged, and which could be considered in future research. This section

---

⁴ The present study did not incorporate any planned or structured interviews with teachers about their attitudes towards providing CF within their own classroom context. The claim that content teachers might be hesitant to use CF is based on informal conversations between the researcher and the teachers.

⁵ Taken together, these observations might lead one to conclude that the schools participating in the present research may not have fully implemented the language sensitive approach. A comparable observation was already done by Swain and Carrol (1987, as cited in Lyster, 2007), who noticed that, in immersion classrooms, form and function were kept surprisingly distinct even though the pedagogical goal was to teach language through content.
consecutively discusses (1) the specific context of the work at hand, (2) its scope (i.e. L2 accuracy development), and (3) some methodological issues.

6.5.1 Research context

It has been pointed out, that it is important for research to acknowledge its contextualization (e.g. Manchón, 2009; Ortega, 2009). In the case of the present work, this is of particular significance, because its setting is very different from the contexts other written CF studies were conducted in. On the one hand, this could be considered a strength of the investigations at hand; several scholars have called for research within diverse contexts (e.g. Manchón, 2009; Ortega, 2009; Révész, 2007). On the other hand, it is not self-evident that findings across contexts are readily comparable, or that findings from one setting can automatically be transferred to any other context.

The present research context differs from that of other recent studies into the efficacy of written CF (e.g. Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009; Bitchener & Knoch, 2010a; Ellis et al., 2008; Sheen, 2007; Sheen, 2010b; Truscott & Hsu, 2008) with respect to the age of the learners, the language under investigation, and learners’ L2 proficiency level. Whereas in other CF studies, participants were generally adults with an intermediate level of L2 English proficiency, the population in the present research consisted of adolescent, highly proficient L2 learners of Dutch.

Additionally, while earlier CF work tested CF efficacy within instructed SLA (i.e. at university level), the educational setting under investigation in the present book could be characterized as an (early) immersion context. As was explained in Chapters 1, 3, and 4, the secondary schools participating in the research project adopted a language sensitive approach to content teaching (e.g. Van Eerde & Hajer, 2005). This pedagogical approach opts to integrate content and language instruction, in order to cater for the special needs of L2 learners. Where the L2 is the main object of study in instructed SLA settings, the first and foremost focus of the present instructional context is on content (e.g. biology).

One conclusion from the present work, that might be dependent on the context it was derived from, is that only direct CF leads to grammatical accuracy development. It needs to be acknowledged that the participants in the studies at hand received a limited

---

6 The research contexts of some of the early CF studies (e.g. Kepner, 1991; Lalande, 1982) differed from the present setting on an additional factor. They investigated the effect of CF within foreign language (FL) instruction rather than in L2 contexts.

7 Participants had an average of 10 years experience in the target language; most of them started learning Dutch in school at age four and were 14 years old at the time of data collection.
amount of explicit grammar instruction. They could therefore be expected to have low levels of meta-linguistic knowledge and awareness. As previously suggested in Chapter 4, it might well be that learners had too little knowledge about the L2 to be able to really and fully benefit from indirect CF, and that another research context would have led to different conclusions on the differential value of direct and indirect CF. It would thus be interesting to replicate the studies presented in this thesis in other pedagogical settings and with other populations, to test the generalizability of the present findings.

6.5.2 Research scope
As was explained in Chapter 2, the focus of the studies presented in this dissertation was exclusively on the effects of written CF on learners’ accuracy development (i.e. the outcome of the language learning process). As a result, the present work only enabled further understanding of CF’s potential in yielding a learning effect, in terms of accuracy gains. It did not provide any direct insights, however, into the SLA processes – such as noticing (the gap), hypothesis testing, problem-solving – that may underlie CF efficacy (e.g. Adams, 2003; Santos et al., 2010; Storch & Wigglesworth, 2010; Swain & Lapkin, 2002; Qi & Lapkin, 2001), and/or the affective factors that may explain learners’ (lack of) feedback uptake (e.g. Bruton, 2010; Goldstein, 2006; Hyland, 1998; Storch & Wigglesworth, 2010; Swain & Lapkin, 2003).

In particular, the findings from the qualitative study presented in Chapter 5 would have been more illuminating when an attempt would have been made to tap into learner internal processes and/or learners’ perceptions of, attitudes towards, and levels of motivation associated with the different CF treatments. In order to gain a full understanding of the role of CF in SLA, future research that triangulates product data, process data (e.g. think-aloud protocols, stimulated recall protocols), and socio-cultural data (e.g. retrospective interviews, questionnaires) would thus be warranted.

6.5.3 Research methodology
Even though the present empirical work tried to address the design-related problems of earlier written CF studies (e.g. Chandler, 2003; Kepner 1991; Polio et al., 1998; Semke, 1984; see Chapter 3 for a review), the present methodological set-up still had its flaws.

To begin with, the studies reported in Chapters 3, 4, and 5 did not include immediate post-tests out of practical considerations and time constraints. Instead, the first post-test was only administered one week after the treatment session (i.e. the moment of feedback provision). Inclusion of an immediate post-test would have enabled the comparison of
immediate and delayed CF effects, and could have provided insights into the proportion of initial accuracy gains (i.e. as evident in the performance on an immediate post-test) that are retained over a longer period of time (i.e. as evident in the performance on a post-test and delayed post-test).

The second shortcoming also relates to the number of post-tests used in the present work. Although the study presented in Chapter 4 already proved that the accuracy gains brought about by comprehensive written CF were retained over a one-month period, it would have been interesting to further investigate the durability of comprehensive CF effects, by means of adding (a) more delayed post-test(s).

A final drawback that needs to be acknowledged, concerns the operationalization of learners’ meta-linguistic awareness in Chapter 4. Learners’ educational level was presumed to be indicative of their level of meta-linguistic awareness. The rationale behind this assumed relationship was twofold (cf. Chapter 2, section 2.6.3). First, it was suggested that pupils from different educational strands could be anticipated to vary in (among other things) their level of meta-linguistic awareness, because a considerable part of the assessment that is used in the Netherlands to place pupils within different levels of secondary education concerns learners’ (meta-)linguistic skills (Cito, 2010). Secondly, it was noted that the Dutch national framework of reference for language skills sets different goals for pupils in different levels of secondary education with respect to language proficiency and meta-linguistic knowledge (Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2008). However, the study reported in Chapter 4 did not test if pupils with a higher level of education indeed disposed of a greater awareness of the L2 system. Without empirical evidence on the validity of its operationalization, any firm claims on the mediating effect of learners’ meta-linguistic awareness on CF efficacy would be unwarranted. Hence, the conclusion that Chapter 4’s results do not constitute evidence for the hypothesis that indirect CF is more helpful for learners with higher levels of meta-linguistic awareness (e.g. Ferris, 2004; Hyland and Hyland, 2006), is a very tentative one.

6.6 Concluding remarks

Whereas the value of written CF for L2 acquisition has been heavily contested (e.g. Truscott, 1996; 1999; 2007), the studies presented in this thesis (as well as other recent CF studies) provide robust evidence on the efficacy of error correction in L2 writing. The fact that the accuracy improvement brought about by written CF was shown to be durable,
rebuts Truscott’s (1996) claim that correction can only lead to a superficial and transient type of L2 knowledge. I conclude that, by offering learners opportunities to notice the gaps in their developing L2 systems, test interlanguage hypotheses, and engage in metalinguistic reflection, written CF has the ability to foster SLA and to lead to accuracy development. The present empirical work thus advances the theoretical understanding of the language learning potential of written CF, and shows that comprehensive CF is a useful pedagogical tool.